

Università degli Studi di Roma “La Sapienza”  
Dipartimento di Informatica e Sistemistica “A. Ruberti”

**Proff. Gianni Di Pillo and Laura Palagi**

**Note per il corso di  
OTTIMIZZAZIONE (a.a. 2007-08)**

Dipartimento di Informatica e Sistemistica “A. Ruberti”, via Buonarroti 12 - 00185 Roma.

**E-mail:** dipillo@dis.uniroma1.it

**E-mail:** palagi@dis.uniroma1.it

**URL:** <http://www.dis.uniroma1.it/~or/gestionale/ottimizzazione>



# Contents

<b>1</b>	<b>Introduzione</b>	<b>1</b>
1.1	Forme standard e prime definizioni . . . . .	1
<b>2</b>	<b>Modelli di Programmazione matematica</b>	<b>5</b>
<b>3</b>	<b>Condizioni di esistenza della soluzione</b>	<b>9</b>
3.1	Esercizi sulle condizioni di esistenza . . . . .	10
<b>4</b>	<b>Forme e funzioni quadratiche</b>	<b>12</b>
<b>5</b>	<b>Soluzione grafica di problemi di ottimizzazione</b>	<b>15</b>
<b>6</b>	<b>Problemi di ottimizzazione convessi</b>	<b>17</b>
<b>7</b>	<b>Le condizioni di ottimalità.</b>	<b>21</b>
7.1	Generalità . . . . .	21
7.2	Direzioni di discesa . . . . .	21
7.3	Il caso non vincolato . . . . .	24
7.4	Esempi sulle condizioni di ottimo non vincolato . . . . .	26
7.5	Il caso vincolato: preliminari . . . . .	36
7.6	Il caso vincolato: vincoli di uguaglianza lineari . . . . .	37
7.7	Il caso vincolato: vincoli di uguaglianza non lineari . . . . .	43
7.8	Analisi di sensibilità per vincoli di uguaglianza . . . . .	48
7.9	Il caso vincolato: vincoli di disuguaglianza . . . . .	52
7.10	Il caso vincolato: vincoli di disuguaglianza e di uguaglianza . . . . .	55
7.11	Il caso convesso . . . . .	57
<b>8</b>	<b>La dualità nella Programmazione Lineare</b>	<b>59</b>
<b>9</b>	<b>Generalità sugli algoritmi di ottimizzazione</b>	<b>64</b>
<b>10</b>	<b>Algoritmi per l'ottimizzazione non vincolata</b>	<b>68</b>
10.1	Introduzione . . . . .	68
10.2	Metodi di ricerca unidimensionale . . . . .	71
10.2.1	Metodi di ricerca esatta . . . . .	72
10.2.2	Ricerca di linea esatta nel caso quadratico . . . . .	72
10.2.3	Metodi di ricerca inesatta: metodo di Armijo . . . . .	73
10.3	Il metodo del gradiente . . . . .	74
10.4	Il metodo del gradiente per funzioni quadratiche . . . . .	75
10.5	Il metodo del gradiente con ricerca di linea . . . . .	81
10.6	Il metodo del gradiente coniugato . . . . .	83
10.7	Il metodo delle direzioni coniugate per funzioni quadratiche . . . . .	83
10.8	Metodo del gradiente coniugato per funzioni quadratiche . . . . .	87
10.9	Il metodo del gradiente coniugato nel caso non quadratico . . . . .	90
10.10	Il metodo di Newton . . . . .	91
10.11	Il metodo di Newton con ricerca di linea . . . . .	93
<b>A</b>	<b>Richiami sulle norme</b>	<b>97</b>

<b>B</b>	<b>Richiami sulla differenziazione in <math>R^n</math></b>	<b>99</b>
B.1	Derivate del primo ordine di una funzione reale . . . . .	99
B.2	Differenziazione di un vettore di funzioni . . . . .	101
B.3	Derivate del secondo ordine di una funzione reale . . . . .	102
B.4	Teorema della media e formula di Taylor . . . . .	103

# 1 Introduzione

## 1.1 Forme standard e prime definizioni

Come messo in luce nel corso di *Ricerca Operativa*, molti problemi dell'Ingegneria Gestionale assumono la forma di un problema di *Ottimizzazione*. Un problema di Ottimizzazione consiste nel determinare il valore di un vettore di *variabili di decisione*  $x \in \mathbb{R}^n$  che minimizza una *funzione obiettivo*  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , quando  $x$  è vincolato ad appartenere ad un *insieme ammissibile*  $\mathcal{F} \subseteq \mathbb{R}^n$ ; cioè consiste nel problema:

$$\min_{x \in \mathcal{F}} f(x). \quad (1)$$

Osserviamo subito che un problema di massimo si può sempre ricondurre a un problema di minimo, cambiando di segno la funzione obiettivo. Infatti, i punti di massimo (ove esistano) del problema

$$\max_{x \in \mathcal{F}} f(x)$$

coincidono con i punti di minimo del problema

$$\min_{x \in \mathcal{F}} -f(x)$$

e risulta:  $\max_{x \in \mathcal{F}} f(x) = -\min_{x \in \mathcal{F}} (-f(x))$ .

In base a tale osservazione ci si può riferire esclusivamente, senza perdita di generalità, a problemi di minimizzazione.

Riportiamo la prima definizione utile.

**Definizione 1 (Punto di minimo globale)** *Un punto  $x^* \in \mathcal{F}$  si dice punto di minimo globale (o assoluto) di  $f$  su  $\mathcal{F}$  se risulta:*

$$f(x^*) \leq f(x), \quad \text{per ogni } x \in \mathcal{F},$$

*e, in tal caso, si dice che  $f(x^*)$  è il minimo (o il valore minimo) globale di  $f$  su  $\mathcal{F}$ , ossia*

$$f(x^*) = \min_{x \in \mathcal{F}} f(x).$$

*Si dice che  $x^* \in \mathcal{F}$  è un punto di minimo globale stretto di  $f$  su  $\mathcal{F}$  se risulta:*

$$f(x^*) < f(x), \quad \text{per ogni } x \in \mathcal{F}, x \neq x^*. \quad \square$$

È opportuno mettere in evidenza che, assegnati  $\mathcal{F}$  e  $f : \mathcal{F} \rightarrow \mathbb{R}$  potrebbero anche non esistere soluzioni ottime. Una prima possibilità è che l'insieme ammissibile  $\mathcal{F}$  sia vuoto; in tal caso non esistono *punti ammissibili* e di conseguenza non esistono soluzioni ottime.

Se  $\mathcal{F}$  è non vuoto, possono verificarsi, nel caso generale, le situazioni seguenti:

- la funzione obiettivo è illimitata inferiormente su  $\mathcal{F}$  ossia:

$$\inf_{x \in \mathcal{F}} f(x) = -\infty$$

e, in tal caso, non esiste un valore minimo di  $f$  su  $\mathcal{F}$ ;

- la funzione obiettivo è limitata inferiormente su  $\mathcal{F}$  ossia:

$$\inf_{x \in \mathcal{F}} f(x) > -\infty,$$

ma tuttavia *non esistono punti di minimo globale* di  $f$  su  $\mathcal{F}$ ;

- esistono punti di minimo globale di  $f$  su  $\mathcal{F}$ ; in tal caso la funzione obiettivo è necessariamente limitata inferiormente su  $\mathcal{F}$  e si ha

$$\inf_{x \in \mathcal{F}} f(x) = \min_{x \in \mathcal{F}} f(x).$$

Solo nell'ultimo caso, ovviamente, ci si può porre il problema della ricerca di una soluzione ottima.

“Risolvere” un problema di ottimizzazione può quindi significare, in pratica:

- stabilire se l'insieme ammissibile è non vuoto, oppure concludere che non esistono soluzioni ammissibili;
- stabilire se esistono soluzioni ottime, oppure dimostrare che il problema non ammette soluzioni ottime;
- determinare (eventualmente in modo approssimato) una soluzione ottima.

Due casi sono di particolare interesse: - l'insieme ammissibile  $\mathcal{F}$  coincide con  $\mathbb{R}^n$ , cosicchè il Problema (1) diviene:

$$\min_{x \in \mathbb{R}^n} f(x); \tag{2}$$

in questo caso si dice che il Problema (1) è *non vincolato*. Più in generale, il Problema (1) è non vincolato se  $\mathcal{F}$  è un insieme aperto in  $\mathbb{R}^n$ . Le condizioni di ottimalità per il Problema (2) considerate nel seguito, non cambiano se  $\mathcal{F}$  è un aperto; quindi in queste note, per semplicità assumiamo che  $\mathcal{F} = \mathbb{R}^n$ .

- l'insieme ammissibile è descritto da *vincoli di disuguaglianza* e/o *vincoli di uguaglianza* sulle variabili di decisione:

$$\mathcal{F} = \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = 1, \dots, p; h_j(x) = 0, j = 1, \dots, m\};$$

in questo caso il Problema (1) diviene:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & g(x) \leq 0 \\ & h(x) = 0, \end{aligned} \tag{3}$$

con  $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$  e  $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ . In questo caso diciamo che il Problema (1) è *vincolato*.

I problemi di Ottimizzazione vengono anche chiamati problemi di *Programmazione Matematica* quando si vuole mettere l'enfasi sui metodi risolutivi dei problemi stessi.

Nel corso di Ricerca Operativa è stato esemplificato il caso della Programmazione Lineare, che corrisponde al caso in cui tutte le funzioni  $f, g_i, i = 1, \dots, p, h_j, j = 1, \dots, m$  del Problema (3) sono combinazioni lineari delle variabili di decisione, sono cioè funzioni della forma:

$$v(x) = c_1x_1 + c_2x_2 + \dots c_nx_n,$$

ove  $c_i, i = 1, \dots, n$  sono coefficienti.

Notiamo che un problema di Programmazione Lineare è necessariamente vincolato, poiché altrimenti si tratterebbe sempre di un problema illimitato.

Si dice che il Problema (1) è un problema di *Programmazione Nonlineare* (PNL) se almeno una, tra le funzioni  $f, g_i, i = 1, \dots, p, h_j, j = 1, \dots, m$ , del Problema (2) o del Problema (3) risulta essere non lineare, rispetto ad almeno una delle componenti del vettore delle variabili di decisione  $x$ .

Per il problema vincolato (3) si assume usualmente che il numero  $m$  di vincoli di uguaglianza non sia maggiore del numero  $n$  di variabili di decisione, cioè si assume  $m \leq n$ . Altrimenti, dovendo le  $n$  variabili soddisfare  $m$  equazioni, l'insieme ammissibile potrebbe risultare vuoto, a meno che alcuni vincoli non siano tra di loro dipendenti, e quindi ridondanti. Una limitazione analoga non vale invece per i vincoli di disuguaglianza.

A volte, tra i vincoli di disuguaglianza, si mettono in esplicita evidenza i *vincoli semplici* sulle variabili, vincoli che esprimono limitazioni sul valore minimo  $m_i$  e massimo  $M_i$  che una variabile  $x_i$  può assumere. In questo caso, il Problema (3) diviene:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & g(x) \leq 0 \\ & h(x) = 0 \\ & m \leq x \leq M. \end{aligned} \tag{4}$$

Se nel Problema (4) la variabile  $x_i$  non è limitata inferiormente (superiormente), si assume per convenzione che  $m_i = -\infty$  ( $M_i = +\infty$ ).

In generale, in problemi di PNL la ricerca di soluzioni globali può risultare difficile, e può avere interesse anche la ricerca di soluzioni di tipo "locale". Per poter definire il concetto di "punto di minimo locale" occorre introdurre il concetto di *intorno sferico aperto*  $\mathcal{S}$  di un punto. In particolare dato  $x^*$ , un intorno sferico aperto  $\mathcal{S}$  di centro  $x^*$  e raggio  $\rho > 0$  è definito come

$$\mathcal{S}(x^*, \rho) = \{x \in \mathbb{R}^n : \|x - x^*\| < \rho\}.$$

Possiamo allora introdurre la seguente definizione

**Definizione 2 (Punto di minimo locale)** *Un punto  $x^* \in \mathcal{F}$  si dice punto di minimo locale (o relativo) di  $f$  su  $\mathcal{F}$  se esiste un intorno  $\mathcal{S}(x^*, \rho)$  di  $x^*$  tale che:*

$$f(x^*) \leq f(x), \quad \text{per ogni } x \in \mathcal{F} \cap \mathcal{S}(x^*, \rho),$$

*e, in tal caso, si dice che  $f(x^*)$  è un minimo locale di  $f$  su  $\mathcal{F}$ .*

*Si dice che  $x^* \in \mathcal{F}$  è un punto di minimo locale stretto di  $f$  su  $\mathcal{F}$  se esiste un intorno  $\mathcal{S}(x^*, \rho)$  di  $x^*$  tale che:*

$$f(x^*) < f(x), \quad \text{per ogni } x \in \mathcal{F} \cap \mathcal{S}(x^*, \rho), \quad x \neq x^*. \quad \square$$

È immediato rendersi conto del fatto che *un punto di minimo globale è anche un punto di minimo locale*. Notiamo anche che nella definizione precedente l'intorno  $\mathcal{S}(x^*, \rho)$  preso in considerazione *non è necessariamente tutto contenuto* in  $\mathcal{F}$ . Nel caso particolare in cui  $\mathcal{F}$  abbia un interno non vuoto ed esista  $\mathcal{S}(x^*, \rho) \subseteq \mathcal{F}$ , tale che  $f(x^*) \leq f(x)$  per ogni  $x \in \mathcal{S}(x^*, \rho)$ , diremo che  $x^*$  è un punto di minimo locale *non vincolato* di  $f$  su  $\mathcal{F}$ .

In queste note assumiamo che le funzioni del problema  $f, g, h$  siano almeno continuamente differenziabili in  $\mathbb{R}^n$ ; quando richiesto, assumiamo che siano almeno due volte continuamente differenziabili.

Nel seguito diremo che un problema di ottimizzazione è scritto in *forma standard* se è scritto nella forma (2) nel caso non vincolato, nella forma (3) nel caso vincolato.

### **Domande**

- 1) Come si mette in forma standard un problema non vincolato in cui la funzione obiettivo è da massimizzare?
- 2) Ogni problema con vincoli di disuguaglianza e uguaglianza può essere trasformato in un problema con soli vincoli di disuguaglianza. Come?
- 3) Ogni problema con vincoli di disuguaglianza e uguaglianza può essere trasformato in un problema con soli vincoli di uguaglianza. Come?

## 2 Modelli di Programmazione matematica

### Esempio 1 (Discriminazione del prezzo)

Consideriamo un monopolista che operi su due mercati distinti (ad es. nazionale ed estero) ciascuno con una diversa funzione di domanda. Indichiamo con  $x_i$  l'offerta sul mercato  $i = 1, 2$  e con  $P_i = f_i(x_i)$  la funzione di domanda inversa sul mercato  $i$ . Il ricavo derivante dalla vendita di  $x_i$  unità di prodotto sul mercato  $i$  è  $x_i f_i(x_i)$ . Supponiamo inoltre che il costo unitario di produzione dipenda solo dal prodotto finale e non dal mercato e sia pari a  $c$ . Il problema consiste nel massimizzare il profitto del monopolista. La funzione ricavo è

$$x_1 f_1(x_1) + x_2 f_2(x_2),$$

mentre il costo è  $c(x_1 + x_2)$ . Il profitto totale sarà quindi

$$f(x) = x_1 f_1(x_1) + x_2 f_2(x_2) - c(x_1 + x_2).$$

Naturalmente possono essere presenti vincoli definiti dal processo di produzione del bene e dal mercato su cui il bene viene immesso. Questi vincoli sono specifici del processo di produzione e del mercato e non entriamo qui nel dettaglio. Li indicheremo semplicemente con  $x \in S$ . Posto  $n = 2$  e  $x = (x_1, \dots, x_n)'$  si ha l'insieme ammissibile:

$$\mathcal{F} = \{x \in S : x \geq 0\}.$$

Più in generale nel caso di  $n$  mercati distinti, il problema di ottimizzazione corrispondente è

$$\max_{x \in \mathcal{F}} \sum_{i=1}^n x_i f_i(x_i) - c \sum_{i=1}^n x_i$$

Molto spesso le funzioni  $f_i(x_i)$  hanno un andamento lineare del tipo

$$f_i(x_i) = a_i - m_i x_i \quad \text{con } m_i > 0.$$

La funzione ricavo risulta essere quindi una funzione quadratica del tipo

$$\sum_{i=1}^n x_i (a_i - m_i x_i) = -\frac{1}{2} x' Q x + a' x$$

con  $Q$  matrice diagonale definita positiva con elementi diagonali  $2m_i > 0$  e  $a = (a_1 \dots, a_n)'$ . Il problema di discriminazione del prezzo diventa un problema di programmazione quadratica del tipo

$$\max_{x \in \mathcal{F}} -\frac{1}{2} x' Q x + (a + c)' x$$

### Esempio 1 (Problemi di approssimazione ai Minimi quadrati)

Supponiamo che siano noti  $n$  punti del piano  $(x_i, y_i)$  con  $i = 1, \dots, n$  che possono corrispondere ai valori una funzione continua  $\phi : R \rightarrow R$  ottenuti per via sperimentali o con misurazioni.

Si vuole approssimare la funzione  $y = \phi(x)$  per mezzo di un polinomio di primo grado (ovvero una retta) del tipo  $y = mx + q$ .

Si definiscono gli errori

$$e_i(x_i) = y_i - (mx_i + q), \quad i = 1, \dots, n,$$

e si può considerare il problema di ottimizzazione non vincolata, (noto anche come problema di *curve fitting*)

$$\min \|e(x)\|^2 = \sum_{i=1}^n e_i(x_i)^2 = \sum_{i=1}^n (mx_i + q - y_i)^2$$

Si osservi che la funzione obiettivo è quadratica.

Più in generale, si può considerare il problema

$$\begin{aligned} \min \quad & \|e(x)\|^\alpha, \\ e \in & R^n \end{aligned}$$

in cui  $\|\cdot\|$  è una norma su  $R^n$  e  $\alpha > 0$ . I casi più comuni sono quelli in cui si richiede di minimizzare una norma  $\ell_p$  con  $p \geq 1$ , o, equivalentemente, la  $p$ -ma potenza di una norma  $\ell_p$ :

$$f(x) = \sum_{i=1}^m |e_i(x)|^p,$$

oppure la norma  $\ell_\infty$ :

$$f(x) = \max_{1 \leq i \leq m} |e_i(x)|.$$

Problemi differenti si ottengono ovviamente in corrispondenza ad altre scelte delle funzioni approssimanti che, nel caso più generale, possono dipendere in modo non lineare dai parametri incogniti.

**Esercizio 1** *Un'industria chimica intende utilizzare della lamiera metallica residua, costruendo un serbatoio scoperto da adibire all'immagazzinamento di un prodotto liquido. La lamiera può essere tagliata e saldata a piacere, è disponibile per complessivi  $150\text{m}^2$  e la si vuole utilizzare tutta. Il serbatoio deve essere contenuto in un capannone a pianta quadrata, con lato di  $10\text{m}$ , e con tetto spiovente dall'altezza di  $4.5$  all'altezza di  $3\text{m}$ . Per semplicità di progetto, si assume che il serbatoio abbia la forma di un prisma retto, con base quadrata.*

*Determiniamo le dimensioni del serbatoio, in modo da massimizzare il volume del liquido che vi può essere contenuto.*

**Soluzione.** Le variabili di decisione sono  $x_1$  la misura del lato di base del serbatoio e  $x_2$  la misura dell'altezza. Il volume del serbatoio è

$$V = A_b \cdot h = x_1^2 x_2.$$

Per quanto riguarda i vincoli abbiamo:

*vincoli di disponibilità:* deve essere usata esattamente una quantità di lamiera pari a  $150$  mq. Quindi, poichè il serbatoio è scoperto la quantità di lamiera necessaria è pari all'area di base  $A_b$  e alle 4 superfici laterali. Quindi  $A_b + 4A_l = 150$  che corrisponde a

$$x_1^2 + 4x_1 x_2 = 150.$$

*Vincoli di spazio:* il serbatoio deve essere collocato nel capannone, quindi

$$x_1 \leq 10;$$

per quanto riguarda  $x_2$ , poiché l'altezza del capannone è variabile da 4.5 a 3 metri e il lato è 10 m. abbiamo

$$x_2 \leq -0.15x_1 + 4.5.$$

*Vincoli di non negatività:* si tratta di lunghezze e quindi

$$x_1 \geq 0, \quad x_2 \geq 0.$$

Complessivamente possiamo scrivere:

$$\begin{aligned} \max \quad & x_1^2 x_2 \\ & x_1^2 + 4x_1 x_2 = 150 \\ & x_1 \leq 10 \\ & x_2 + 0.15x_1 \leq 4.5 \\ & x_1 \geq 0, \quad x_2 \geq 0 \end{aligned}$$

Osserviamo che l'insieme ammissibile è compatto. Infatti  $0 \leq x_1 \leq 10$  e  $0 \leq x_2 \leq 4.5 - 0.15x_1 \leq 4.5$ . La funzione è continua ovunque, quindi esiste un punto di minimo globale.

**Esercizio 2** *Un'industria dolciaria si rifornisce di zucchero acquistandolo in tre diversi paesi produttori, che indicheremo con  $A, B, C$ . I prezzi di acquisto dello zucchero nei tre paesi sono diversi, e, in ciascuno dei tre paesi, il prezzo subisce delle variazioni aleatorie, dovute alle variazioni dei cambi, alla maggiore o minore produzione stagionale, alle variazioni salariali ecc.. Per tener conto di queste variazioni aleatorie i prezzi per tonnellata sono caratterizzati mediante i valori medi  $p_A, p_B, p_C$ , e le varianze  $\sigma^2_A, \sigma^2_B, \sigma^2_C$ . Ovviamente se i prezzi non subissero variazioni aleatorie, converrebbe acquistare tutta la materia prima del paese che offre il prezzo minore. In presenza delle incertezze sui prezzi, l'industria deve fare riferimento, nella propria programmazione degli acquisti, ad un prezzo medio stimato  $p_M$ , e acquista lo zucchero nei tre paesi secondo proporzioni tali da realizzare questo prezzo medio stimato. Indicando con  $x_1, x_2, x_3$  le variabili che rappresentano, per ogni tonnellata di zucchero, le frazioni acquistate rispettivamente nei paesi  $A, B, C$  deve risultare:*

$$\begin{aligned} x_1 p_A + x_2 p_B + x_3 p_C &= p_M \\ x_1 + x_2 + x_3 &= 1; \end{aligned}$$

*il primo vincolo infatti esprime il fatto che il prezzo medio stimato dello zucchero sia pari  $p_M$  per ogni tonnellata, e il secondo vincolo esprime il fatto che le variabili di decisioni sono frazioni di una quantità unitaria (la tonnellata).*

*Come obiettivo della programmazione degli acquisti, l'industria assume quello di minimizzare il rischio che il prezzo medio effettivo per tonnellata risultante dagli acquisti effettuati secondo le frazioni  $x_1, x_2, x_3$ , differisca da quello stimato  $p_M$ ; ciò per rendere il più possibile certe, in termini di costi di produzione, le conseguenze della programmazione degli acquisti. Una ovvia misura di questo rischio è data dalla varianza del prezzo medio  $\sigma^2_M$ , quantità data in questo caso dalla espressione*

$$\sigma^2_M = x_1^2 \sigma^2_A + x_2^2 \sigma^2_B + x_3^2 \sigma^2_C.$$

Assumiamo che, nell'unità monetaria adottata risulti:

$$\begin{aligned} p_A &= 4, & p_B &= 5.5 & p_C &= 6 \\ \sigma_A^2 &= 1, & \sigma_B^2 &= 0.8, & \sigma_C^2 &= 0.5 \end{aligned}$$

e che l'industria abbia programmato sulla base di un prezzo medio  $P_M = 5$ .

**Soluzione.** Si tratta di un problema di programmazione quadratica

$$\begin{aligned} \min \quad & x_1^2 + 0.8x_2^2 + 0.5x_3^2 \\ & x_1 + x_2 + x_3 = 1 \\ & 4x_1 + 5.5x_2 + 6x_3 = 5 \end{aligned}$$

**Esercizio 3** Una compagnia petrolifera si rifornisce di greggio in tre città portuali, che indicheremo con  $A, B, C$ . Il porto  $B$  è ubicato 300Km a est e 400Km a nord del porto  $A$ , e il porto  $C$  è ubicato 400Km ad est e 100Km a sud del porto  $B$ . La compagnia intende costruire una nuova raffineria per il greggio, e intende localizzare la nuova raffineria in modo da minimizzare la quantità totale di tubi occorrenti per collegare la raffineria ai porti. Determiniamo la formulazione di questo problema di localizzazione.

Supponiamo inoltre che non sia possibile situare la raffineria né a sud del porto  $A$ , né entro un raggio di 360 Km. dallo stesso.

**Soluzione.** Scegliamo un sistema di riferimento con il porto  $A$  nell'origine. Il porto  $B$  ha quindi coordinate  $(300,400)$ , mentre il porto  $C$   $(700,300)$ . La raffineria si trova nella posizione incognita  $(x_1, x_2)$  e l'obiettivo è minimizzare la distanza, cioè

$$\min (x_1^2 + x_2^2)^{1/2} + ((x_1 - 300)^2 + (x_2 - 400)^2)^{1/2} + ((x_1 - 700)^2 + (x_2 - 300)^2)^{1/2}$$

Inoltre abbiamo i vincoli sulle coordinate della raffineria:

$$\begin{aligned} x_2 &\geq 0 \\ x_1^2 + x_2^2 &\geq 360^2 \end{aligned}$$

### 3 Condizioni di esistenza della soluzione

Stabilire l'esistenza di soluzioni di un problema di ottimo, a partire dalla caratterizzazione analitica della funzione obiettivo e dell'insieme ammissibile può essere, in generale, difficile. Una semplice condizione *sufficiente* (ma non necessaria) per l'esistenza di un punto di minimo globale in un problema di ottimo in cui lo spazio delle variabili sia  $\mathbb{R}^n$  è quella riportata nel teorema di Teorema di Weierstrass.

**Teorema 1 (Teorema di Weierstrass)** *Sia  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una funzione continua e sia  $\mathcal{F} \subset \mathbb{R}^n$  un insieme compatto. Allora esiste un punto di minimo globale di  $f$  in  $\mathcal{F}$ .*

Il risultato stabilito nel Teorema 1 si applica in modo diretto solamente a problemi *vincolati* in cui l'insieme ammissibile  $\mathcal{F} \subseteq \mathbb{R}^n$  è compatto ovvero è chiuso e limitato.

Osserviamo che il teorema dà una condizione *sufficiente* di esistenza, ma non *necessaria*: ad esempio, il semplice problema

$$\begin{aligned} \min_{x \in \mathbb{R}} \quad & x^2 \\ & x \geq 0, \end{aligned}$$

ha la soluzione globale  $x^* = 0$ , anche se l'insieme ammissibile, costituito da tutto il semiasse  $x \geq 0$  è chiuso, ma non limitato; analogamente, il problema:

$$\begin{aligned} \min_{x \in \mathbb{R}} \quad & x^2 \\ & -1 < x < 1, \end{aligned}$$

ha la soluzione globale  $x^* = 0$  anche se l'insieme ammissibile è costituito dall'intervallo  $(-1, 1)$ , limitato ma non chiuso.

Notiamo che se la funzione  $g_i(x)$  è continua, l'insieme dei punti che soddisfano il vincolo  $g_i(x) \leq 0$  è un insieme chiuso (e quindi, se  $h_j(x)$  è continua, anche l'insieme dei punti che soddisfano il vincolo  $h_j(x) = 0$  è chiuso). Pertanto nei casi che prenderemo in considerazione la chiusura di  $S$  è sempre assicurata. Pertanto, l'applicazione del Teorema di Weierstrass al Problema (3) si riconduce alla verifica che l'insieme ammissibile  $\mathcal{F}$  sia limitato. Ciò è sicuramente vero se tutte le variabili  $x_i$  sono limitate da valori finiti sia inferiormente che superiormente, come in pratica accade spesso nei problemi dell'Ingegneria Gestionale.

Nel caso non vincolato tuttavia, in cui  $\mathcal{F} = \mathbb{R}^n$  non è possibile applicare direttamente il teorema di Weierstrass. Allo scopo di definire condizioni sufficienti di esistenza per il problema (2), si definisce l'*insieme di livello*  $\mathcal{L}_{x^0}$  della funzione obiettivo  $f(x)$  relativo ad un punto  $x^0$  come l'insieme dei punti in cui la funzione ha valore minore o eguale a quello di  $f(x^0)$ :

$$\mathcal{L}_{x^0} = \{x \in \mathbb{R}^n : f(x) \leq f(x^0)\}.$$

Possiamo concludere che il Problema (2) ha sicuramente una soluzione globale se per qualche  $x^0$  l'insieme di livello  $\mathcal{L}_{x^0}$  risulta compatto: infatti, in questo caso,  $f(x)$  ha un minimo globale  $x^*$  in  $\mathcal{L}_{x^0}$ , ed al di fuori di  $\mathcal{L}_{x^0}$  risulta sicuramente  $f(x) > f(x^0) \geq f(x^*)$ .

Una funzione  $f(x)$  si dice *radialmente illimitata*, o, anche, *coerciva*, se gode della proprietà di tendere a  $+\infty$  quando ci si allontana dall'origine in  $\mathbb{R}^n$ : cioè se risulta:

$$\lim_{\|x\| \rightarrow \infty} f(x) = +\infty.$$

Si può dimostrare che, se  $f(x)$  è radialmente illimitata i suoi insiemi di livello sono compatti per ogni valore di  $x^0$ , e quindi una funzione radialmente illimitata ha un punto di minimo globale in  $\mathbb{R}^n$ .

La coercività di  $f$  è quindi una condizione solo sufficiente di esistenza di un punto di minimo globale di  $f$  come illustrato in Figura 1 in cui è rappresentata una funzione non coerciva che ammette un minimo globale.

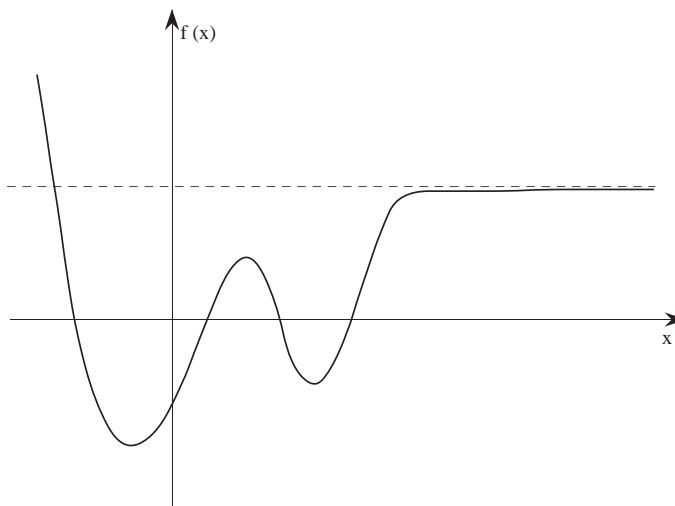


Figure 1: Esempio di funzione non coerciva con un minimo globale.

Più in generale, nell'applicazione del Teorema 1, possiamo limitarci a prendere in considerazione i punti di  $\mathcal{F}$  nell'insieme  $\mathcal{L}_{x^0}$ , ossia i punti di  $\mathcal{F}$  in cui la funzione obiettivo abbia valore non superiore a quello assunto nel punto  $x^0$ . Vale quindi il risultato seguente.

**Teorema 2 (Condizione sufficiente di esistenza)** *Sia  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una funzione continua e sia  $\mathcal{F} \subset \mathbb{R}^n$  un insieme chiuso e non vuoto. Supponiamo che esista  $x^0 \in \mathcal{F}$  tale che l'insieme di livello*

$$\mathcal{L}_{x^0} = \{x \in \mathcal{F} : f(x) \leq f(x^0)\}$$

*sia limitato. Allora esiste un punto di minimo globale di  $f$  su  $\mathcal{F}$ .*

**Dim.** Poiché  $\mathcal{F}$  è chiuso e  $f$  è continua, l'insieme  $\mathcal{L}_{x^0}$  è chiuso e quindi, per l'ipotesi fatta, è anche compatto. Per il Teorema 1 esiste allora un punto di minimo  $x^*$  di  $f$  su  $\mathcal{L}_{x^0}$  a cui corrisponde il valore minimo  $f(x^*) \leq f(x^0)$ . D'altra parte, se  $x \in \mathcal{F}$  non appartiene a  $\mathcal{L}_{x^0}$  ciò implica, per definizione di insieme di livello che sia  $f(x) > f(x^0) \geq f(x^*)$  e ciò implica che  $x^*$  è un punto di minimo globale su tutto  $\mathcal{F}$ .  $\square$

### Domande

1) Si supponga che nel Problema (3) l'insieme ammissibile sia chiuso ma non limitato, e la funzione obiettivo sia radialmente illimitata. Cosa si può dire sull'esistenza della soluzione?

### 3.1 Esercizi sulle condizioni di esistenza

Studiamo l'esistenza del minimo di alcune funzioni.

**Esempio 2** *La semplice funzione quadratica:*

$$f(x) = x_1^2 + x_2^2 + \dots + x_n^2 = \|x\|^2$$

è evidentemente radialmente illimitata, e il suo minimo globale è nel punto  $x_i^* = 0, i = 1, \dots, n$ .

**Esempio 3** *Sia data la funzione*

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

*Studiare l'esistenza di punti di minimo.*

**Soluzione.** La funzione è coerciva. Infatti possiamo scrivere

$$f(x_1, x_2) = (x_1^4 + x_2^4) \left( 1 - \frac{3x_1x_2}{x_1^4 + x_2^4} \right)$$

risulta, comunque presa una successione  $x^k$ , tale che  $\|x^k\| \rightarrow \infty$

$$\lim_{\|x\| \rightarrow \infty} \frac{3x_1x_2}{x_1^4 + x_2^4} = 0$$

Si ha quindi, quindi  $\lim_{\|x\| \rightarrow \infty} f(x_1, x_2) = \infty$  e la funzione è coerciva. Poiché la funzione è anche continua ammette almeno un minimo globale.

**Esempio 4** *Sia data la funzione*

$$f(x_1, x_2) = x_1^3 - 12x_1x_2 + 8x_2^3.$$

*Dire se esiste un minimo globale.*

**Soluzione.** La funzione non è coerciva. Infatti lungo la direzione definita da  $x_2 = \bar{x}_2 = \text{cost.}$  risulta

$$\lim_{\substack{\|x\| \rightarrow \infty \\ x_1 \rightarrow -\infty}} f(x_1, \bar{x}_2) = -\infty.$$

In particolare, poiché esiste una direzione lungo cui la funzione è illimitata inferiormente, possiamo concludere anche che non esiste un punto di minimo globale.

## 4 Forme e funzioni quadratiche

Data una matrice  $A$  quadrata e simmetrica di dimensione  $(n \times n)$ , con elementi  $a_{ij} = a_{ji}$ ,  $i, j = 1, \dots, n$ , si definisce *forma quadratica* associata alla matrice  $A$  la funzione:

$$x^T Ax = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j. \quad (5)$$

Come sarà chiaro nel seguito, le forme quadratiche hanno un ruolo importante nell'Ottimizzazione. Una *funzione quadratica* è una funzione del tipo

$$q(x) = \frac{1}{2} x^T Ax + c^T x$$

Una forma quadratica è quindi una particolare funzione quadratica in cui il termine lineare è identicamente nullo ( $c = 0$ ).

In questo paragrafo consideriamo le proprietà di forme e funzioni quadratiche di cui si fa uso corrente.

La forma quadratica  $x^T Ax$  e, corrispondentemente, la matrice  $A$  associata alla forma quadratica si dicono:

- *definita positiva* ( $A \succ 0$ ), se risulta  $x^T Ax > 0 \forall x \in \mathbb{R}^n, x \neq 0$ ;
- *semidefinita positiva* ( $A \succeq 0$ ), se risulta  $x^T Ax \geq 0 \forall x \in \mathbb{R}^n$ ;
- *indefinita* se per qualche  $x \in \mathbb{R}^n$  risulta  $x^T Ax > 0$ , e per altri  $x \in \mathbb{R}^n$  risulta  $x^T Ax < 0$ .

La forma quadratica  $x^T Ax$  si dice (semi)definita negativa se  $-x^T Ax$  è (semi)definita positiva.

Corrispondentemente, la matrice  $A$  associata alla forma quadratica si dice (semi)definita negativa se la matrice  $-A$  è (semi)definita positiva. Nel seguito ci riferiremo in modo indifferente a forme quadratiche o alla matrice  $A$  associata alla forma quadratica.

Per verificare se una matrice  $A$  è definita positiva, si può utilizzare un semplice test.

**Criterio 1** *Siano  $A_k, k = 1 \dots, n$  gli  $n$  minori principali della matrice  $A$ , detti sottomatrici principali di nord-ovest, cioè le  $n$  sottomatrici con elementi  $a_{ij}, i, j = 1, \dots, k$ , ottenute da  $A$  eliminando le ultime  $n - k$  righe e colonne. Denotato con  $\det A_k$  il determinante di  $A_k$ , risulta che:*

- $A$  è definita positiva se, e solo se,  $\det A_k > 0$ , per  $k = 1, \dots, n$ .

Si osserva che se  $A$  è semidefinita positiva allora risulta  $\det A_k \geq 0$ , ma non è vero in generale il viceversa. Basta prendere come esempio la matrice

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 2a_{22} \end{pmatrix} \quad (6)$$

con  $a_{22} < 0$  e la forma quadratica associata

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$$

in cui  $a_{11} = a_{12} = 0$ . I minori principali  $A_1$  e  $A_2$  hanno determinante nullo, e quindi soddisfano il criterio  $\det A_k \geq 0$ , ma  $q(x)$  è semidefinita negativa !!!

Per verificare se una matrice  $A$  è semidefinita positiva si deve applicare un criterio molto più oneroso riportato di seguito.

**Criterio 2** Siano  $A_{I_k}$  le sottomatrici di  $A$  di dimensione  $k \in [1, n]$  con elementi  $a_{ij}$  con  $i, j \in I_k$ , dette minori principali di ordine  $k$  di  $A$ .  $A$  è semidefinita positiva se, e solo se, tutti i minori principali per  $k = 1, \dots, n$  sono non negativi, ovvero, denotato con  $\det A_{I_k}$  il determinante di  $A_{I_k}$ ,

-  $A$  è semidefinita positiva se, e solo se,  $\det A_{I_k} \geq 0$  per ogni sottoinsieme di righe/colonne  $I_k \subseteq \{1, \dots, n\}$  di cardinalità  $k$  con  $1 \leq k \leq n$ .

Ad esempio in una matrice di dimensione  $n = 3$ , le matrici  $A_{I_k}$  per  $k = 1, 2, 3$  sono quelle corrispondenti ai seguenti insiemi di indici:  $\{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 3\}, \{1, 2, 3\}$ . Tra questi minori, ci sono ovviamente i minori di Nord-ovest che corrispondono agli insiemi  $\{1\}, \{1, 2\}, \{1, 2, 3\}$ .

Applicando il criterio alla matrice (6), si verifica facilmente che non è soddisfatto.

Per analizzare se una matrice  $A$  è definita/semidefinita negativa si possono applicare i criteri precedenti alla matrice  $-A$  associata alla forma quadratica.

Un altro test per verificare il segno di una forma quadratica consiste nel determinare gli *autovalori* della matrice  $A$ , cioè i valori  $\alpha_i$ ,  $i = 1, \dots, n$  che risolvono l'equazione di grado  $n$ :

$$\det(A - \alpha I) = 0,$$

ove  $I$  è la matrice identica di ordine  $n$ . Se la matrice  $A$  è simmetrica, si ha che gli autovalori sono tutti reali. Risulta che:

- $A$  è definita positiva se, e solo se,  $\alpha_i > 0$ , per  $i = 1, \dots, n$ ;
- $A$  è semidefinita positiva se, e solo se,  $\alpha_i \geq 0$  per  $i = 1, \dots, n$ ;
- $A$  è indefinita, altrimenti.

È evidente che il test basato sui minori è di più semplice impiego di quello basato sugli autovalori: infatti calcolare determinanti, fino all'ordine  $n$ , è più semplice che risolvere un'equazione di grado  $n$ .

Nel seguito indicheremo  $\alpha_{\min}(A)$  e  $\alpha_{\max}(A)$  rispettivamente il più piccolo e il più grande autovalore di una matrice  $A$ .

Si verifica facilmente che il gradiente e l'Hessiano della forma quadratica sono dati rispettivamente da:

$$\text{gradiente} = Ax, \quad \text{hessiano} = A.$$

Ricordiamo inoltre che per ogni forma quadratica del tipo  $x^T Ax$ , vale la seguente caratterizzazione

$$\alpha_{\min}(A)\|x\|^2 \leq x^T Ax \leq \alpha_{\max}(A)\|x\|^2. \quad (7)$$

Abbiamo il seguente risultato.

**Teorema 3** Sia data una funzione quadratica

$$q(x) = \frac{1}{2}x^T Ax + c^T x$$

Condizione necessaria e sufficiente affinché  $q(x)$  sia coerciva è che  $A$  sia definita positiva.

**Dim.** Dimostriamo che se  $A$  è definita positiva allora  $q(x)$  è coerciva. Utilizzando la (7) e il fatto che  $c^T x \geq -|c^T x| \geq -\|c\|\|x\|$ , possiamo scrivere

$$\frac{1}{2}x^T Ax + c^T x \geq \frac{1}{2}\alpha_{\min}(A)\|x\|^2 - |c^T x| \geq \frac{1}{2}\alpha_{\min}(A)\|x\|^2 - \|c\|\|x\| = \|x\| \left( \frac{1}{2}\alpha_{\min}(A)\|x\| - \|c\| \right)$$

Per  $\|x\|$  sufficientemente grande, poiché  $\alpha_{\min}(A) > 0$  risulta  $\frac{1}{2}\alpha_{\min}(A)\|x\| - \|c\| > 0$ , e quindi  $\lim_{\|x\| \rightarrow \infty} q(x) = \infty$ .

**(non in programma)** Dimostriamo ora che se  $q(x)$  è coerciva allora  $A$  è definita positiva. Procediamo per assurdo e supponiamo che  $q(x)$  sia coerciva e  $\alpha_{\min}(A) \leq 0$ . Sia  $u$  l'autovettore corrispondente all'autovalore  $\alpha_{\min}(A)$ , cioè

$$Au = \alpha_{\min}(A)u.$$

Si ha

$$q(\lambda u) = \frac{1}{2}\lambda^2 u^T Au + \lambda c^T u = -\frac{1}{2}\lambda^2 |\alpha_{\min}(A)| \|u\|^2 + \lambda c^T u.$$

Se  $\alpha_{\min}(A) = 0$  la funzione diventa  $q(\lambda u) = \lambda c^T u$ , cioè una funzione lineare che non è coerciva.

Se invece  $\alpha_{\min}(A) < 0$ , ricordando che  $c^T x \leq |c^T x| \leq \|c\|\|x\|$ , possiamo scrivere

$$q(\lambda u) \leq -\frac{1}{2}\lambda^2 |\alpha_{\min}(A)| \|u\|^2 + \lambda \|c\| \|u\| = \alpha \|u\| \left( -\frac{1}{2}\lambda |\alpha_{\min}(A)| \|u\| + \|c\| \right)$$

Al limite per  $\lambda \rightarrow \infty$ , la quantità tra parentesi è negativa e risulta  $q(\lambda u) \rightarrow -\infty$ , contraddicendo l'ipotesi di coercività.  $\square$

### Domande

1) Si supponga che la matrice  $A$  sia data dal prodotto di una matrice  $B$  ( $n \times m$ ) per la trasposta di  $B$ :  $A = BB^T$ . Cosa si può dire sulla forma quadratica associata ad  $A$ ?

## 5 Soluzione grafica di problemi di ottimizzazione

In questo paragrafo, consideriamo problemi di ottimizzazione lineare in due variabili e otteniamo la soluzione per via grafica.

**Esempio 2** Sia la funzione obiettivo da minimizzare:

$$f(x) = x_1$$

e sia l'insieme ammissibile  $F$  definito dai vincoli:

$$\begin{aligned} (x_1 - 3)^2 + (x_2 - 2)^2 &= 13 \\ (x_1 - 4)^2 + x_2^2 &\leq 16 \end{aligned}$$

Determinare, se esiste, un punto di minimo.

**Soluzione.**

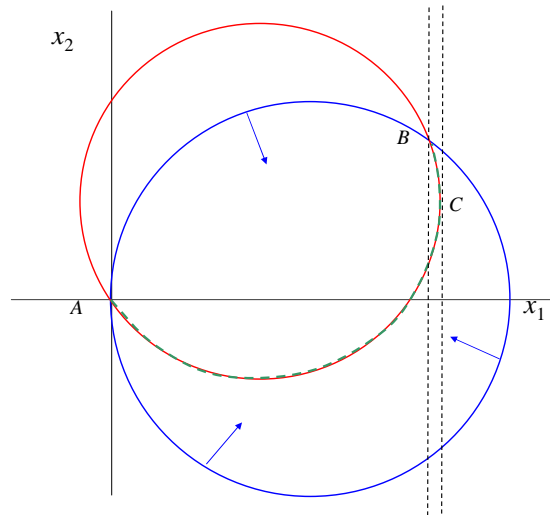


Figure 2: Soluzione grafica in  $\mathbb{R}^2$  dell'Esempio 2.

L'insieme ammissibile  $\mathcal{F}$ , rappresentato in Figura (2), ed è dato dall'arco di circonferenza  $\widehat{ACB}$  (tratteggiato in verde). Poiché esso è compatto e  $f$  è continua, in base al Teorema di Weierstrass, l'esistenza di un minimo globale è assicurata.

Dallo studio del grafico, disegnate le curve di livello  $x_1 = k$  al variare di  $k$ , si determina il punto di minimo globale nel punto  $A = (0,0)^T$  di valore  $f^* = 0$ . Si osserva che il punto  $C$  è un massimo globale, mentre il punto  $B$  risulta essere un minimo locale.

**Esempio 3** Sia dato il problema di ottimizzazione vincolata

$$\begin{aligned} \min \quad & 2x_1 + 3x_2 \\ & 6x_1 + x_2^2 - x_2 \leq 12 \\ & x_1 - 4x_2 \leq -2 \\ & x_1, x_2 \geq 0 \end{aligned}$$

Determinare, se esiste, un punto di minimo globale.

**Esempio 4** Si consideri la funzione obiettivo:

$$f(x) = \frac{1}{2}(x_1^2 + x_2^2)$$

e sia l'insieme ammissibile  $\mathcal{F}$  definito dai vincoli:

$$\begin{aligned}x_1 + x_2 &= 1 \\(x_1 - 1)^2 + x_2^2 &\leq \frac{1}{4}\end{aligned}$$

Determinare se esiste un punto di minimo.

**Soluzione.** L'insieme ammissibile  $F$  é dato dal segmento di retta che ha per estremi i punti  $\tilde{x} = \left(1 - \frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}\right)^T$  e  $\bar{x} = \left(1 + \frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}\right)^T$ .

L'esistenza di una soluzione é garantita dal Teorema di Weierstrass poiché la funzione obiettivo é continua e l'insieme ammissibile é compatto.

Le curve di livello della funzione obiettivo

$$\frac{1}{2}(x_1^2 + x_2^2) = r^2$$

sono circonferenze di centro l'origine e raggio  $r$ .

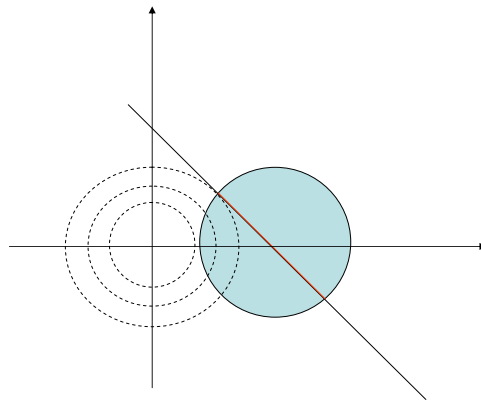


Figure 3: Soluzione grafica in  $\mathbb{R}^2$  dell'Esempio 4.

## 6 Problemi di ottimizzazione convessi

Tra i problemi di Ottimizzazione sono di particolare interesse i cosiddetti problemi *convessi*.

Ricordiamo che un insieme  $\mathcal{C} \subseteq \mathbb{R}^n$  è convesso se, presi comunque due punti  $y, z \in \mathcal{C}$ , risulta che anche  $[y, z] \subseteq \mathcal{C}$ , avendo denotato con  $[y, z]$  il segmento che congiunge  $y$  e  $z$ , segmento dato dai punti  $x$  ottenuti come:

$$x = (1 - \beta)y + \beta z, \quad \beta \in [0, 1].$$

Si verifica facilmente che l'intersezione di un numero finito di insiemi convessi è un insieme convesso.

**Definizione 3 (Funzione (strettamente) convessa)** Una funzione  $v(x)$  si dice convessa su un insieme convesso  $\mathcal{C}$  se, presi comunque due punti  $y, z \in \mathcal{C}$  risulta che:

$$v((1 - \beta)y + \beta z) \leq (1 - \beta)v(y) + \beta v(z), \quad \beta \in [0, 1]. \quad (8)$$

La funzione  $v(x)$  si dice strettamente convessa se, per  $y, z \in \mathcal{C}$ ,  $y \neq z$ , risulta

$$v((1 - \beta)y + \beta z) < (1 - \beta)v(y) + \beta v(z), \quad \beta \in (0, 1).$$

□

Una funzione  $v(x)$  si dice (strettamente) concava su un insieme convesso  $\mathcal{C}$  se la funzione  $-v(x)$  è (strettamente) convessa su  $\mathcal{C}$ .

Nella (8)  $y, z, v(y), v(z)$  sono dati, e  $\beta$  varia tra 0 e 1. Se mettiamo in esplicita evidenza la dipendenza da  $\beta$ , introducendo la funzione

$$\phi(\beta) = v((1 - \beta)y + \beta z)$$

otteniamo per una funzione strettamente convessa che:

$$\phi(\beta) < (1 - \beta)\phi(0) + \beta\phi(1) \quad \beta \in (0, 1)$$

Quest'ultima relazione mette in evidenza che, se si rappresenta graficamente nel piano  $(\beta, \phi)$  la funzione  $\phi(\beta)$ , il grafico della funzione nell'intervallo  $(0, 1)$  si trova al di sotto del segmento, detto *secante*, che congiunge i punti  $(0, \phi(0))$  e  $(1, \phi(1))$  e coincide solo negli estremi del segmento. Si può concludere che una funzione strettamente convessa è caratterizzata dalla proprietà di avere il grafico sempre al di sotto di ogni sua secante.

A partire dalla definizione di funzione (strettamente) convessa, è possibile dedurre due importanti proprietà che riguardano le derivate, rispettivamente prime e seconde.

**Teorema 4** Una funzione  $v(x)$  è convessa su un insieme convesso  $\mathcal{C}$  se, e solo se, per ogni  $y, z \in \mathcal{C}$  risulta:

$$v(z) \geq v(y) + \nabla v(y)^T(z - y);$$

$v(x)$  è strettamente convessa su  $\mathcal{C}$  se, e solo se, per ogni  $y, z \in \mathcal{C}$ ,  $y \neq z$ , risulta:

$$v(z) > v(y) + \nabla v(y)^T(z - y).$$

In termini di funzione  $\phi(\beta)$ , la precedente disuguaglianza si riscrive come:

$$\phi(1) > \phi(0) + d\phi(0)/d\beta$$

Pertanto, utilizzando la rappresentazione grafica nel piano  $(\beta, \phi)$ , si può concludere che una funzione strettamente convessa ha il grafico sempre al di sopra di ogni sua tangente.

**Teorema 5** Una funzione  $v(x)$  è convessa su un insieme convesso  $\mathcal{C}$  se, e solo se, per ogni  $x \in \mathcal{C}$  risulta:

$$\frac{1}{2}y^T \nabla^2 v(x)y \geq 0, \quad \text{per ogni } y \in \mathbb{R}^n;$$

inoltre, se risulta

$$\frac{1}{2}y^T \nabla^2 v(x)y > 0, \quad \text{per ogni } y \in \mathbb{R}^n,$$

$v(x)$  è strettamente convessa su  $\mathcal{C}$ .

Si noti che l'ultima condizione è solo sufficiente: ad esempio, la funzione  $v(x) = x^4$  è strettamente convessa in  $\mathcal{C} = \mathbb{R}^n$ , ma non soddisfa la condizione, in quanto in  $x = 0$  la derivata seconda si annulla. Nel caso di funzioni quadratiche le condizioni del teorema diventano invece necessarie e sufficienti, ovvero si ha

**Teorema 6** Una funzione quadratica del tipo  $q(x) = \frac{1}{2}x^T Qx + c^T x$  è convessa su un insieme convesso  $\mathcal{C}$  se e solo se risulta  $Q$  semidefinita positiva. Inoltre,  $q(x)$  è strettamente convessa su  $\mathcal{C}$  se e solo se risulta  $Q$  definita positiva.

**Definizione 4 (Problema (strettamente) convesso)** Si dice che il Problema (1) è un problema di ottimizzazione convesso se l'insieme ammissibile  $\mathcal{F}$  è un insieme convesso e la funzione obiettivo  $f(x)$  è una funzione convessa su  $\mathcal{F}$ . Se la funzione obiettivo è strettamente convessa su  $\mathcal{F}$ , il problema si dice strettamente convesso.

I problemi di ottimizzazione convessi sono di particolare importanza per due motivi. Il primo è che la grande maggioranza dei problemi di ottimizzazione che si incontrano nella pratica sono convessi (vedi la Programmazione Lineare). Il secondo è che la convessità induce alcune proprietà che semplificano l'analisi e la soluzione di un problema convesso.

La prima di tali proprietà (altre ne vedremo nel seguito) consiste nel fatto che:

**Teorema 7** Un problema di ottimizzazione convesso o non ha soluzione, o ha solo soluzioni globali; non può avere soluzioni esclusivamente locali.

**Dimostrazione.** La dimostrazione è per assurdo. Ammettiamo che  $x^*$  sia una soluzione locale, ma non globale, di un problema convesso  $\min_{x \in \mathcal{C}} f(x)$ . Allora esisterà un altro punto  $z \in \mathcal{C}$  tale che  $f(z) < f(x^*)$ . La costruzione utilizzata nella dimostrazione è illustrata schematicamente nella figura 4.

Consideriamo il segmento  $[x^*, z]$ : per la convessità di  $f$ , si ha:

$$f((1 - \beta)x^* + \beta z) \leq (1 - \beta)f(x^*) + \beta f(z) = f(x^*) + \beta(f(z) - f(x^*)), \text{ per ogni } \beta \in [0, 1].$$

Il termine  $\beta(f(z) - f(x^*))$  risulta  $< 0$ , per ogni  $\beta \in (0, 1]$ , e si annulla solo per  $\beta = 0$ . Quindi, poichè in ogni punto  $x \in (x^*, z]$  risulta  $f(x) < f(x^*)$ , non esiste nessun intorno di raggio  $\rho > 0$  in cui  $x^*$  può soddisfare la definizione di minimo locale.  $\square$

Una seconda proprietà notevole è espressa dalla seguente proposizione:

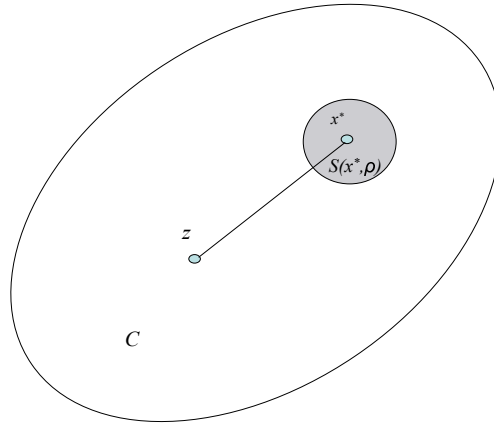


Figure 4: Rappresentazione della costruzione geometrica usata nelle dimostrazione.

**Teorema 8** *In un problema di ottimizzazione strettamente convesso la soluzione globale, se esiste, è unica.*

**Dimostrazione.** Anche questa dimostrazione è per assurdo, e si lascia per esercizio. Sia  $f(x)$  strettamente convessa. Si ammetta, per assurdo, che  $x^*$  e  $x^{**}$ , con  $x^* \neq x^{**}$ , siano due soluzioni globali di un problema convesso, e se ne traggano le conseguenze.◁

Riconoscere che un problema di ottimizzazione è convesso fornisce quindi importanti informazioni qualitative sulle sue soluzioni. Per riconoscere che un problema è convesso dobbiamo verificare che  $\mathcal{F}$  è un insieme convesso, e che  $f(x)$  è convessa su  $\mathcal{F}$ , il che non è sempre facile. Ci aiuta la seguente proposizione, di facile utilizzo, e che è verificata molto spesso nella pratica. La proposizione fornisce una condizione sufficiente affinché un problema sia convesso.

**Teorema 9** *Si assuma che nel Problema (3) la funzione obiettivo  $f(x)$  sia una funzione convessa in  $\mathbb{R}^n$ , che i vincoli di disuguaglianza siano dati da funzioni  $g_i(x)$  convesse in  $\mathbb{R}^n$ , e che i vincoli di uguaglianza siano dati da funzioni affini del tipo  $a_j^T x - b_j$ . Allora il Problema (3) è convesso.*

**Dimostrazione.** Facciamo dapprima vedere che, nelle ipotesi poste, l'insieme ammissibile risulta convesso. Osserviamo che l'insieme ammissibile è esprimibile come:

$$\bigcap_{i=1}^p \{x \in \mathbb{R}^n : g_i(x) \leq 0\} \cap \bigcap_{j=1}^m \{x \in \mathbb{R}^n : a_j^T x - b_j = 0\}.$$

Si tratta dell'intersezione di un numero finito di insiemi. Quindi è sufficiente dimostrare che ciascuno di questi insiemi è convesso.

Consideriamo il generico insieme  $\{x \in \mathbb{R}^n : g_i(x) \leq 0\}$ . Presi due punti  $y$  e  $z$  che soddisfano il vincolo di disuguaglianza  $i$ -esimo, si ha:

$$g_i((1 - \beta)y + \beta z) \leq (1 - \beta)g_i(y) + \beta g_i(z) \leq 0,$$

ove la prima disuguaglianza segue dalla convessità, la seconda segue dal fatto che  $y$  e  $z$  sono ammissibili. Quindi, poichè in tutti i punti  $w$  del segmento  $[y, z]$  risulta  $g_i(w) \leq 0$ , il segmento  $[y, z] \subseteq \{x \in \mathbb{R}^n : g_i(x) \leq 0\}$ , cioè l'insieme dei punti che soddisfano il vincolo è convesso.

Per il vincolo di uguaglianza  $a_j^T x - b_j = 0$ , presi due punti  $y$  e  $z$  che lo soddisfano, si ha

$$a_j^T((1 - \beta)y + \beta z) - b_j = (1 - \beta)(a_j^T y - b_j) + \beta(a_j^T z - b_j) = 0.$$

Quindi, poichè in tutti i punti  $x$  del segmento  $[y, z]$  risulta  $a_j^T x - b_j = 0$ , l'insieme dei punti che soddisfano il vincolo è convesso.

Poichè l'intersezione di un numero finito di insiemi convessi è un insieme convesso, possiamo concludere che, nelle ipotesi poste, l'insieme ammissibile del Problema (3) risulta convesso.

Per concludere la dimostrazione, basta osservare che una funzione  $f(x)$  convessa in  $\mathbb{R}^n$  è convessa anche su ogni sottoinsieme di  $\mathbb{R}^n$ .  $\square$

**Esempio 5** *Dato il problema*

$$\begin{aligned} \max \quad & x_2 \\ & x_2 - x_1^3 \geq 0 \\ & x_1 + x_2 \leq 1 \\ & x_1, x_2 \geq 0 \end{aligned}$$

*Dire se è convesso. Utilizzare la condizione sufficiente espressa dal Teorema 9 e anche la rappresentazione grafica.*

### Domande

- 1) Sia  $f(x)$  convessa in  $\mathbb{R}^n$ . I suoi insiemi di livello sono insiemi convessi?
- 2) Siano  $v_1(x), v_2(x)$  funzioni convesse su un insieme convesso  $\mathcal{C}$ , e siano  $c_1 \geq 0, c_2 \geq 0$  due coefficienti. La funzione  $v(x) = c_1 v_1(x) + c_2 v_2(x)$  è convessa su  $\mathcal{C}$ ? E se uno dei due coefficienti è negativo?
- 3) Un problema di Programmazione Lineare è convesso?
- 4) Un problema di Programmazione Lineare è strettamente convesso?
- 5) Sia  $h_j(x)$  una funzione convessa. Si può affermare che i punti che soddisfano il vincolo  $h_j(x) = 0$  è un insieme convesso?
- 6) Si fornisca un esempio di problema di ottimizzazione convesso per cui non vale la Proposizione (9).

## 7 Le condizioni di ottimalità.

### 7.1 Generalità

Una soluzione locale  $x^*$  di un problema di ottimizzazione deve soddisfare una *condizione necessaria di ottimalità* (CNO). Ad esempio, per il problema non vincolato

$$\min_{x \in \mathbb{R}} f(x), \quad (9)$$

la CNO consiste nel fatto che la derivata della funzione  $f$  si deve annullare in  $x^*$ :  $df(x^*)/dx = 0$ . Ricordiamo però sempre che i punti che soddisfano una CNO per un problema non sono necessariamente soluzioni del problema stesso: ad esempio i punti che annullano la derivata di  $f$  nel Problema (9) possono essere punti di massimo, anziché di minimo. Comunque, la definizione di CNO è di fondamentale importanza nell'Ottimizzazione, poichè detto  $\Omega$  l'insieme dei punti che soddisfa la CNO, e risultando evidentemente  $\Omega \subseteq \mathcal{F}$ , ci si può limitare a cercare la soluzione nell'insieme  $\Omega$  anziché in tutto  $\mathcal{F}$ , che di solito è molto più grande di  $\Omega$ . Per un punto ammissibile generico  $x \in \mathcal{F}$ , la CNO fornisce un *certificato di ottimalità*, nel senso che se  $x \notin \Omega$ ,  $x$  non può essere soluzione del problema, mentre può esserlo (ma può anche non esserlo) se  $x \in \Omega$ .

Se poi un punto  $x^*$  soddisfa una *condizione sufficiente di ottimalità* (CSO) per un problema di ottimizzazione, si può affermare che  $x^*$  è una soluzione locale del problema stesso. Ad esempio, per il Problema (9), una condizione sufficiente di ottimalità è che in  $x^*$  si annulli la derivata prima e sia positiva la derivata seconda:  $df(x^*)/dx = 0$ ,  $d^2f(x^*)/dx^2 > 0$ . Ricordiamo però che un punto  $x^*$  che è soluzione di un problema di ottimizzazione può non soddisfare la CSO per il problema stesso: ad esempio, il problema

$$\min_{x \in \mathbb{R}} f(x) = x^4,$$

ha come soluzione  $x^* = 0$ , anche se la derivata seconda di  $x^4$  in  $x^* = 0$  si annulla, anziché essere positiva. Pertanto, se per un problema di ottimizzazione si è trovato un punto di  $\Omega$  che non soddisfa la relativa CSO, non si può escludere che il punto trovato sia soluzione del problema.

#### Domande

- 1) Quali condizioni di ottimalità sono soddisfatte per il problema  $\min_{x \in \mathbb{R}} f(x) = x^3$  ?
- 2) Cosa si può intendere per *condizione necessaria e sufficiente di ottimalità* per un problema di ottimizzazione?

### 7.2 Direzioni di discesa

Consideriamo il generico problema di ottimizzazione (1)

$$\min_{x \in \mathcal{F}} f(x),$$

ed introduciamo il concetto di *direzione di discesa* che ci consentirà di fare un primo passo nella definizione delle condizioni di ottimalità.

**Definizione 5 (Direzione di discesa)** *Un vettore  $d \in \mathbb{R}^n$  è una direzione di discesa per la funzione  $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$  nel punto  $x$  se esiste uno scalare  $\alpha^{\max} > 0$  tale che risulti:*

$$f(x + \alpha d) < f(x) \text{ per ogni } \alpha \in (0, \alpha_{\max}^{(1)}]. \quad (10)$$

In sostanza, se  $d$  è una direzione di discesa nel punto  $x$ , spostandosi da questo punto di una quantità  $\alpha$  sufficientemente piccola, si è sicuri di ottenere un decremento della funzione  $f$ . La quantità  $\alpha$  viene chiamata *spostamento* lungo la direzione  $d$ . Naturalmente, se lo spostamento  $\alpha$  supera il limite  $\alpha_{\max}^{(1)}$ , può accadere che risulti  $f(x + \alpha d) > f(x)$ , e quindi si abbia un aumento, anzichè una diminuzione della funzione.

Ad esempio, consideriamo la funzione:

$$f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2, \quad (11)$$

e facciamo vedere che nel punto  $x = (0, 0)$  la direzione  $d = (-1, 0)$  è di discesa. Infatti si ha:

$$f(x) = f(0, 0) = 0;$$

inoltre risulta:

$$x + \alpha d = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\alpha \\ 0 \end{pmatrix},$$

e quindi

$$f(x + \alpha d) = f(-\alpha, 0) = -\alpha + 2\alpha^2 = \alpha(2\alpha - 1).$$

Si verifica allora immediatamente che la condizione (10) è soddisfatta, prendendo ad esempio,  $\alpha_{\max}^{(1)} = 0.4$ ; infatti, per ogni  $\alpha \in (0, 0.4]$  si ha  $f(-\alpha, 0) = \alpha(2\alpha - 1) < f(0, 0) = 0$ . Si verifica inoltre che per  $\alpha > \alpha_{\max}^{(1)} = 0.4$  la condizione  $f(x + \alpha d) < f(x)$  non è necessariamente soddisfatta: se si prende  $\alpha = 0.6$  si ha infatti  $f(-0.6, 0) = 0.12 > f(0, 0) = 0$ .

Nel caso di problemi di minimizzazione con due sole variabili di decisione, la direzione di discesa ha un'immediata interpretazione grafica. Se infatti rappresentiamo nel piano  $(x_1, x_2)$  le linee di livello della funzione  $f$ , osserviamo che, data una direzione di discesa in un punto  $x$  su una linea di livello, spostandoci lungo questa direzione, attraversiamo una zona di linee di livello corrispondenti a valori della funzione decrescenti rispetto a quello assunto in  $x$ ; ciò significa muoversi *in discesa* rispetto alle linee di livello della funzione  $f$ .

Ad esempio, nella figura 5 sono rappresentate le linee di livello della funzione (11), e si vede che la direzione  $d$  è di discesa nel punto  $x^a$ .

La figura 5 mette bene in evidenza il carattere *locale* della definizione di direzione di discesa: una direzione  $d$  è di discesa *in un punto*  $x$ , e la discesa si verifica per spostamenti a partire da  $x$  *sufficientemente piccoli*. Dalla figura si vede che la stessa direzione  $d$  non è di discesa (è anzi di salita) nel punto  $x^b$ ; inoltre se lo spostamento dal punto  $x^a$  è eccessivo, si ottiene un aumento, anzichè una diminuzione, della funzione.

Le direzioni di discesa sono caratterizzate dalla seguente condizione:

**Teorema 10** *Condizione sufficiente affinché la direzione  $d$  sia di discesa per la funzione  $f(x)$  nel punto  $x$  è che risulti:*

$$\nabla f(x)^T d < 0. \quad (12)$$

**Dimostrazione.** Basta ricordare che il termine  $\nabla f(x)^T d$  risulta essere la *derivata direzionale* della funzione  $f$  nella direzione  $d$ ; si ha cioè:

$$\lim_{\alpha \rightarrow 0^+} \frac{f(x + \alpha d) - f(x)}{\alpha} = \nabla f(x)^T d.$$

Se al limite il rapporto a secondo membro è  $< 0$ , per  $\alpha$  sufficientemente piccolo deve risultare  $f(x + \alpha d) < f(x)$ .  $\square$

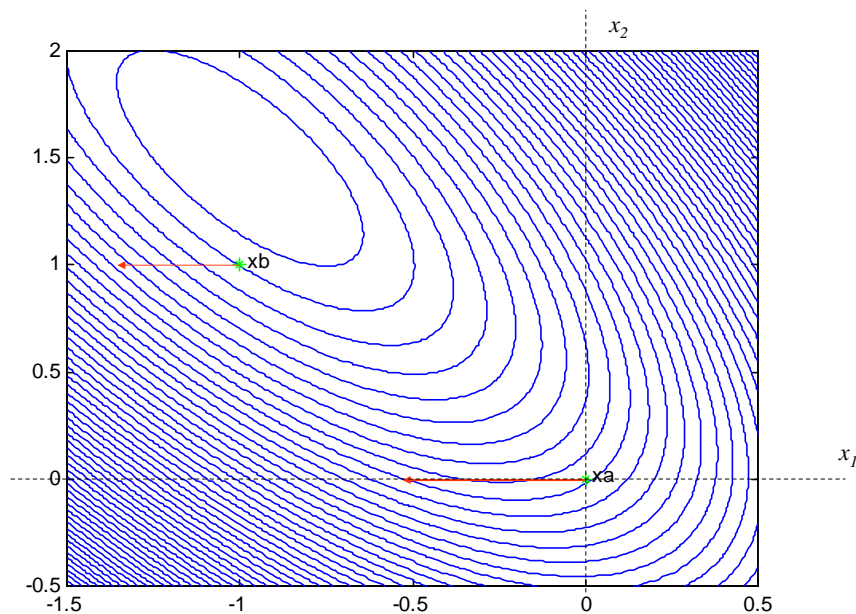


Figure 5: Curve di livello della funzione (11).

Una direzione  $d$  è *di salita* per la funzione  $f(x)$  nel punto  $x$  se  $-d$  è di discesa. Quindi, se risulta  $\nabla f(x)^T d > 0$ , la direzione  $d$  è di salita. Notiamo che se risulta  $\nabla f(x)^T d = 0$  non è possibile dire, basandosi solo sulla conoscenza del gradiente  $\nabla f(x)$ , se  $d$  è di discesa o di salita.

Determinare una direzione che sia di discesa per una funzione  $f(x)$  in un punto  $x$  è molto semplice. Infatti, presa una direzione  $d$  qualsiasi, purchè tale che  $\nabla f(x)^T d \neq 0$ , o risulta  $\nabla f(x)^T d < 0$ , e allora  $d$  è di discesa, o risulta  $\nabla f(x)^T d > 0$ , e allora  $-d$  è di discesa. Ricordando che

$$\nabla f(x)^T d = \|\nabla f(x)\| \|d\| \cos \theta$$

dove  $\theta$  è l'angolo compreso tra  $\nabla f(x)$  e  $d$ , la condizione (12) esprime il fatto che la direzione deve formare un angolo ottuso con la direzione del gradiente. Quindi, in un punto  $x$  tale che  $\nabla f(x) \neq 0$ , la direzione  $d = -\nabla f(x)$ , detta dell'*antigradiente*, è sicuramente di discesa: infatti risulta:

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

Si lascia come esercizio verificare che, in un punto  $x$  tale che  $\nabla f(x) \neq 0$ , ogni direzione del tipo  $d = -A\nabla f(x)$ , ove  $A$  è una matrice  $(n \times n)$  simmetrica e definita positiva, è sicuramente una direzione di discesa.

Ad esempio, nella figura 6 sono rappresentate le linee di livello della funzione (11), evidenziando il vettore gradiente nei due punti  $x^a$  e  $x^b$ .

### Domande

1. Data la funzione (11), il punto  $x = (0, 0)$  e la direzione  $d = (-1, 0)$ , in quale intervallo può variare il valore  $\alpha^{\max}$ ?

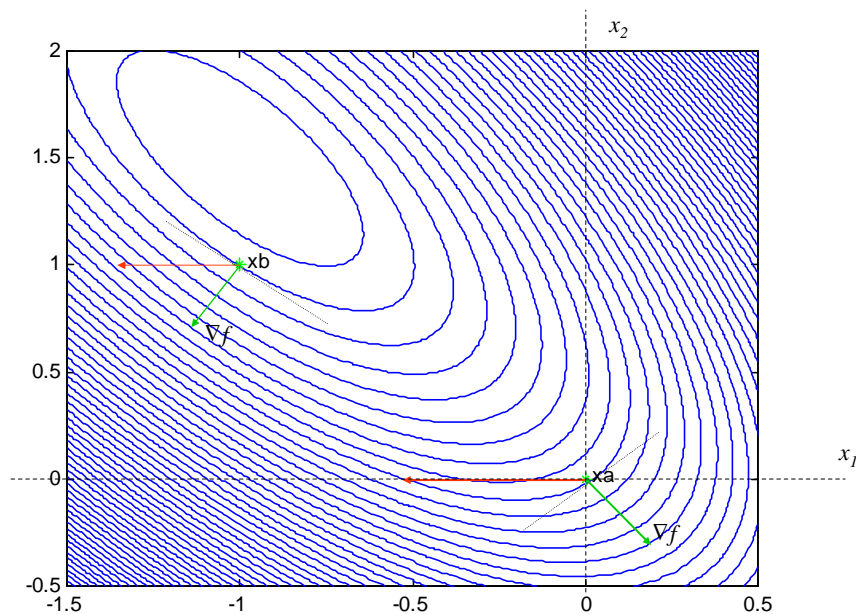


Figure 6: Curve di livello della funzione (11).

- Per una funzione  $f(x)$  due volte continuamente differenziabile, viene definito *curvatura* della funzione nel punto  $x$  e nella direzione  $d$  il termine  $d^T \nabla^2 f(x) d$ . Sai giustificare il fatto che, in un punto  $x^*$  in cui il gradiente della funzione è nullo, una direzione a curvatura negativa è una direzione di discesa? (Suggerimento: utilizza lo sviluppo  $f(x^* + \alpha d) = f(x^*) + \nabla f(x^*)^T d + \frac{\alpha^2}{2} d^T \nabla^2 f(x^*) d + o(\alpha^2)$ ).

### 7.3 Il caso non vincolato

Le condizioni di ottimalità per il problema di ottimizzazione non vincolata (2) sono note dal corso di Analisi Matematica, ma vengono qui riformulate, mettendo in evidenza il ruolo delle direzioni di discesa. È infatti evidente che, se  $x^*$  è un punto di minimo locale per  $f(x)$ , non può esistere una direzione  $d$  che sia di discesa per  $f$  in  $x^*$ ; altrimenti, a partire da  $x^*$  ci si potrebbe muovere lungo  $d$  facendo diminuire la funzione, e  $x^*$  non potrebbe essere un punto di minimo.

Riscriviamo per comodità il problema non vincolato:

$$\min_{x \in \mathbb{R}^n} f(x); \quad (13)$$

Abbiamo innanzi tutto la seguente condizione necessaria di ottimalità:

**Teorema 11** *Condizione necessaria affinché  $x^*$  sia una soluzione locale del Problema (13) è che risulti  $\nabla f(x^*)^T d \geq 0$ , per ogni  $d \in \mathbb{R}^n$ .*

**Dimostrazione.** La dimostrazione segue dal fatto che, se esiste una direzione  $\bar{d}$  tale che  $\nabla f(x^*)^T \bar{d} < 0$ , la direzione  $\bar{d}$  è di discesa per  $f$  in  $x^*$ , e quindi  $x^*$  non può essere un punto di minimo locale per  $f$ .  $\square$

La condizione precedente è di grande importanza concettuale, ma non si presta in pratica a risolvere il Problema (13), in quanto fa intervenire tutti i vettori  $d$  di  $\mathbb{R}^n$ . Una condizione di grande utilità pratica è invece la seguente.

**Teorema 12** *Condizione necessaria affinché  $x^*$  sia una soluzione locale del Problema (13) è che risulti  $\nabla f(x^*) = 0$ .*

**Dimostrazione.** La dimostrazione segue dal fatto che, se  $\nabla f(x^*) \neq 0$ , la direzione  $d = -\nabla f(x^*)$  è di discesa per  $f$  in  $x^*$ .  $\square$

La condizione della proposizione precedente viene detta condizione necessaria *del primo ordine*. I punti che soddisfano  $\nabla f(x) = 0$  vengono detti anche *punti stazionari*.

Se  $f(x)$  è due volte continuamente differenziabile, vale anche la seguente condizione *del secondo ordine*:

**Teorema 13** *Condizione necessaria affinché  $x^*$  sia una soluzione locale del Problema (13) è che risulti  $\nabla f(x^*) = 0$  ed inoltre che la matrice hessiana  $\nabla^2 f(x^*)$  risulti semidefinita positiva.*

**Dimostrazione.** Se  $\nabla^2 f(x^*)$  non è semidefinita positiva, esiste almeno una direzione  $\bar{d}$  tale che  $\bar{d}^T \nabla^2 f(x^*) \bar{d} < 0$ . Tenendo conto del fatto che  $\nabla f(x^*) = 0$ , si ha dalla formula di Taylor:

$$f(x^* + \alpha \bar{d}) = f(x^*) + \frac{\alpha^2}{2} \bar{d}^T \nabla^2 f(x^*) \bar{d} + \dots \text{(termini infinitesimi rispetto ad } \alpha^2 \text{)},$$

e quindi:

$$\lim_{\alpha \rightarrow 0^+} \frac{f(x^* + \alpha \bar{d}) - f(x^*)}{\alpha^2} = \frac{1}{2} \bar{d}^T \nabla^2 f(x^*) \bar{d} < 0.$$

Quest'ultima disuguaglianza mostra che, se  $\nabla f(x^*) = 0$  e  $\nabla^2 f(x^*)$  non è semidefinita positiva, esiste una direzione  $\bar{d}$  di discesa in  $x^*$ , cosicchè  $x^*$  non può essere soluzione del Problema (13).  $\square$

Una condizione sufficiente di ottimalità è data dalla proposizione seguente:

**Teorema 14** *Condizione sufficiente affinché  $x^*$  sia una soluzione locale stretta del Problema (13) è che risulti  $\nabla f(x^*) = 0$  e  $\nabla^2 f(x^*)$  definita positiva.*

**Dimostrazione.** Se  $\nabla^2 f(x^*)$  è definita positiva, per qualsiasi direzione  $d \neq 0$  risulta  $d^T \nabla^2 f(x^*) d > 0$ . Tenendo conto del fatto che  $\nabla f(x^*) = 0$ , si ha dalla formula di Taylor:

$$f(x^* + \alpha d) = f(x^*) + \frac{\alpha^2}{2} d^T \nabla^2 f(x^*) d + \dots \text{(termini infinitesimi rispetto ad } \alpha^2 \text{)},$$

e quindi:

$$\lim_{\alpha \rightarrow 0^+} \frac{f(x^* + \alpha d) - f(x^*)}{\alpha^2} = \frac{1}{2} d^T \nabla^2 f(x^*) d > 0, \text{ per ogni } d \neq 0.$$

Quest'ultima disuguaglianza mostra che, se  $\nabla f(x^*) = 0$  e  $\nabla^2 f(x^*)$  è definita positiva, risulta  $f(x^* + \alpha d) > f(x^*)$  lungo qualsiasi direzione  $d$ , purchè si assuma  $\alpha > 0$  sufficientemente piccolo. Quindi  $x^*$  è una soluzione locale stretta.  $\square$

Nel caso di funzione obiettivo  $f(x)$  nel Problema (13) sia (strettamente) convessa in  $\mathbb{R}^n$ . Si può facilmente dimostrare che la condizione  $\nabla f(x^*) = 0$  è non solo necessaria, ma anche sufficiente affinché  $x^*$  sia una soluzione globale (stretta) del problema.

**Teorema 15 (Condizione sufficiente di minimo globale non vincolato)** Sia  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  continuamente differenziabile su  $\mathbb{R}^n$  e sia  $x^* \in \mathbb{R}^n$ . Si supponga che  $f$  sia convessa. Se  $\nabla f(x^*) = 0$ , allora  $x^*$  è un punto di minimo globale di  $f$  su  $\mathbb{R}^n$ . Inoltre, se  $f$  è strettamente convessa su  $\mathbb{R}^n$ , allora  $x^*$  è l'unico punto di minimo globale di  $f$  su  $\mathbb{R}^n$ .

**Dimostrazione.** Utilizzando la caratterizzazione di funzioni convesse espressa dal teorema 4), si può scrivere

$$f(x) \geq f(x^*) + \nabla f(x^*)^T(x - x^*) \quad \forall x \in \mathbb{R}^n$$

da cui, se  $x^*$  è tale che  $\nabla f(x^*) = 0$  si ottiene  $f(x) \geq f(x^*)$  per ogni  $x \in \mathbb{R}^n$ .  $\square$

I risultati precedenti possono essere specializzati nel caso di funzioni quadratiche.

**Teorema 16** Sia  $f(x) = \frac{1}{2}x'Qx + c'x$ , con  $Q$  simmetrica e  $c \in \mathbb{R}^n$ . Allora:

- (a)  $f(x)$  ammette un unico punto di minimo globale  $x^* = -Q^{-1}c$  se e solo se  $Q$  è definita positiva;
- (b) se  $Q$  è semidefinita positiva ogni punto  $x^*$  tale che  $Qx^* + c = 0$  è un punto di minimo globale di  $f(x)$ ;
- (c)  $f(x)$  ammette un punto di minimo se e solo se  $Q$  è semidefinita positiva ed esiste  $x^*$  tale che  $Qx^* + c = 0$ .

**Dimostrazione.** La dimostrazione segue banalmente osservando che nel caso quadratico risulta  $\nabla f(x) = Qx + c$ ,  $\nabla^2 f(x) = Q$ . Ricordando che  $Q \succ 0$  se e solo se  $f(x)$  è coerciva e strettamente convessa, segue la a). Se  $Q \succeq 0$  risulta  $f(x)$  convessa e quindi la condizione  $\nabla f(x^*) = 0$  è necessaria e sufficiente di minimo globale (ed è dimostrata la b)).

La parte necessaria della c) segue dalle condizioni già date, mentre per la parte sufficiente si osserva che nel caso quadratico risulta

$$f(x^* + \alpha d) = f(x^*) + \alpha(Qx^* + c)^T d + \frac{1}{2}\alpha^2 d^T Q d.$$

Quindi se  $Qx^* + c = 0$  e  $d^T Q d \geq 0$  per ogni  $d \in \mathbb{R}^n$  si ottiene  $f(x^* + \alpha d) \geq f(x^*)$  per ogni  $\alpha \in \mathbb{R}$  e ogni  $d \in \mathbb{R}^n$ .  $\square$

## Domande

1) Quali sono le condizioni di ottimalità per il problema  $\max_{x \in \mathbb{R}} f(x)$ ?

## 7.4 Esempi sulle condizioni di ottimo non vincolato

**Esempio 6** Sia data la funzione dell'Esempio 3

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

Studiare la natura degli eventuali punti stazionari.

**Soluzione.** Abbiamo già verificato nell'Esempio 3 che la funzione ammette almeno un minimo globale che si trova tra i punti che annullano il gradiente. Quindi imponiamo

$$\nabla f = \begin{pmatrix} 4x_1^3 - 3x_2 \\ 4x_2^3 - 3x_1 \end{pmatrix} = 0.$$

Applichiamo le condizioni necessarie del primo ordine; dall'annullamento del gradiente si ottiene il sistema

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

Si ottengono le tre soluzioni:  $A(0,0)$ ,  $B\left(\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2}\right)$ ,  $C\left(-\frac{\sqrt{3}}{2}, -\frac{\sqrt{3}}{2}\right)$ . Tra queste c'è senz'altro il minimo globale.

Le curve di livello della funzione sono nella Figura 7 da cui si individuano i tre punti stazionari.

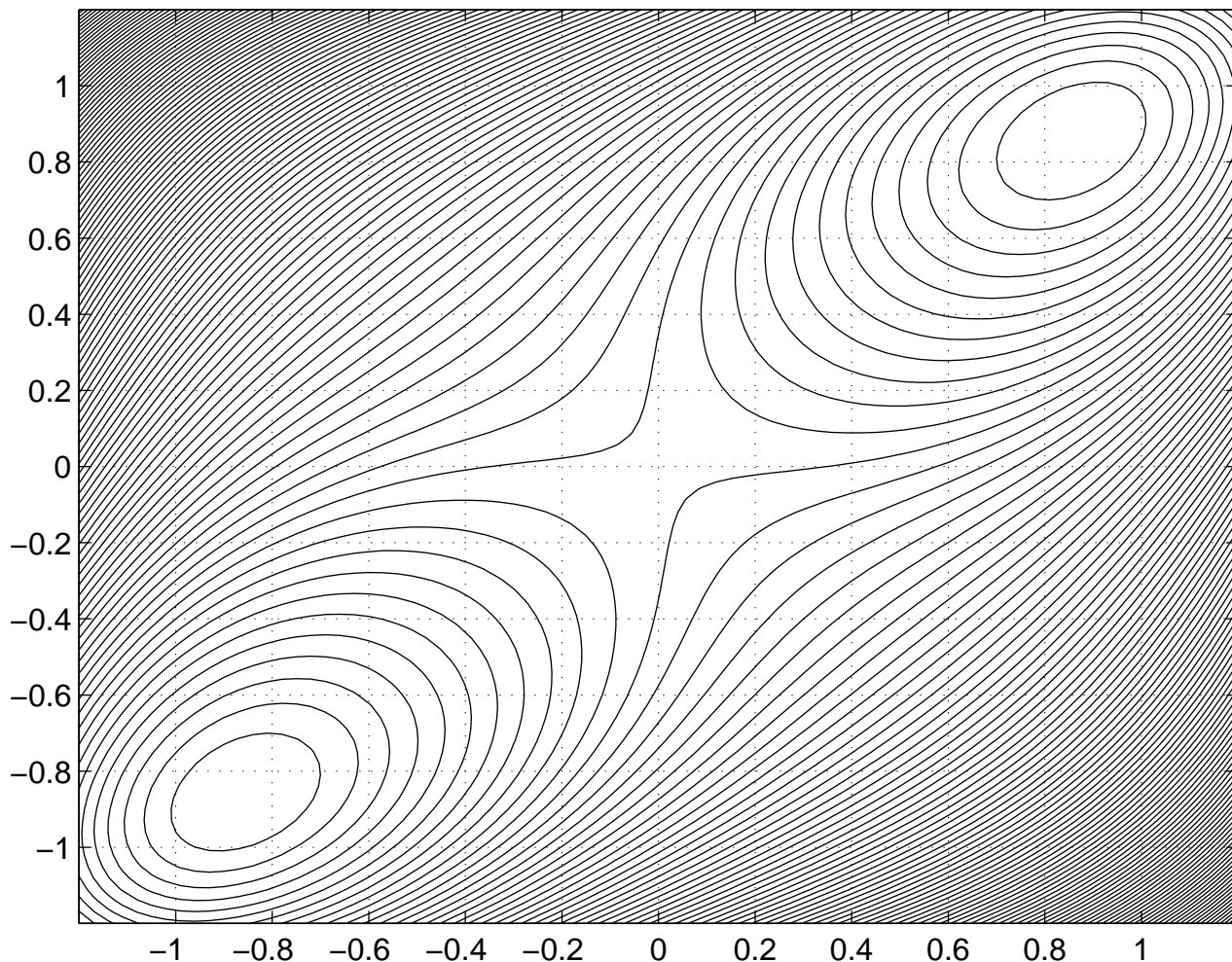


Figure 7: Curve di livello della funzione  $x_1^4 + x_2^4 - 3x_1x_2$ .

Per determinare la natura di  $A, B, C$  utilizziamo le condizioni del secondo ordine. Calcoliamo la matrice hessiana

$$\nabla^2 f = \begin{pmatrix} 12x_1^2 & -3 \\ -3 & 12x_2^2 \end{pmatrix}$$

Risulta nei tre punti:

$$\nabla^2 f(0,0) = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}$$

Quindi in  $A$  la matrice hessiana è indefinita con autovalori  $\lambda_{\min} = -3$ ,  $\lambda_{\max} = 3$ . Si tratta quindi di un punto di sella con valore della funzione  $f(0, 0) = 0$ .

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

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

In questo caso la matrice hessiana è definita positiva con autovalori  $\lambda_{\min} = 2$ ,  $\lambda_{\max} = 4$ . Si tratta quindi di due punti di minimo globale di valore della funzione obiettivo  $f = -\frac{9}{8}$ .

**Esempio 7** *Mostrare che le funzioni*

$$f(x_1, x_2) = x_1^4 + x_2^3$$

e

$$g(x_1, x_2) = x_1^4 + x_2^4,$$

hanno entrambe un punto stazionario in  $(0, 0)$ , con hessiano semidefinito positivo, ma  $(0, 0)$  è un punto di sella per  $f(x_1, x_2)$  e un punto di minimo locale stretto per  $g(x_1, x_2)$ .

**Soluzione.** Abbiamo

$$\nabla f = \begin{pmatrix} 4x_1^3 \\ 3x_2^2 \end{pmatrix} \quad \nabla^2 f = \begin{pmatrix} 12x_1^2 & 0 \\ 0 & 6x_2 \end{pmatrix}$$

e

$$\nabla g = \begin{pmatrix} 4x_1^3 \\ 4x_2^3 \end{pmatrix} \quad \nabla^2 g = \begin{pmatrix} 12x_1^2 & 0 \\ 0 & 12x_2^2 \end{pmatrix}$$

Si verifica facilmente che il punto  $(0, 0)$  è stazionario per  $f$  e  $g$ ; inoltre soddisfa le condizioni necessarie del secondo ordine. In particolare risulta

$$\nabla^2 f(0, 0) = \nabla^2 g(0, 0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \succeq 0.$$

Per determinare la reale natura del punto, consideriamo lo sviluppo delle funzioni  $f, g$  in un punto  $\xi$  nell'intorno del punto  $(0, 0)$ . Ricordiamo che per una generica funzione  $h$  risulta (teorema della media):

$$h(x + \xi) = h(x) + \nabla h(x)^T \xi + \frac{1}{2} \xi^T \nabla^2 h(w) \xi$$

in cui  $w \in (x, x + \xi)$ . Tenendo presente che  $f(0, 0) = g(0, 0) = 0$  e  $\nabla f(0, 0) = \nabla g(0, 0) = (0, 0)^T$  per  $f$  e  $g$  si può scrivere

$$f(\xi) = \frac{1}{2} \xi^T \nabla^2 f(\theta_1 \xi) \xi,$$

$$g(\xi) = \frac{1}{2} \xi^T \nabla^2 g(\theta_2 \xi) \xi.$$

con  $\theta_1 \in (0, 1)$ ,  $\theta_2 \in (0, 1)$ .

Si verifica facilmente che  $\xi^T \nabla^2 g(\theta_2 \xi) \xi > 0$  qualunque sia  $\xi \neq (0, 0)^T$  e quindi  $g(\xi) > g(0, 0) = 0$  per ogni  $\xi$  in un intorno di  $(0, 0)^T$ .

Mentre per  $\xi^T \nabla^2 f(\theta_1 \xi) \xi$  abbiamo che punti  $\xi$  con componente  $\xi_2$  negativa, rendono l'hessiano  $\nabla^2 f$  indefinito e quindi esistono punti appartenenti ad un intorno di  $(0, 0)$  per cui  $f(\xi) > 0$  e altri per cui  $f(\xi) < 0$ . Si tratta quindi di un punto di sella.  $\square$

**Esempio 8** Sia data la funzione

$$q(x) = x_1^2 - 2x_1x_2 + x_2^2$$

Studiare l'esistenza e la natura di eventuali punti estremali.

**Soluzione.** Si tratta di una funzione quadratica. La funzione non é coerciva. Infatti, possiamo scrivere  $q(x) = (x_1 - x_2)^2$ . Quindi se consideriamo come direzione  $x_1 = x_2$ , la funzione é identicamente nulla e quindi

$$\lim_{\|x\| \rightarrow \infty} q(x) = \lim_{\|x\| \rightarrow \infty} (x_1 - x_2)^2 = 0.$$

Osserviamo che, in questo caso, se consideriamo separatamente i limiti

$$\lim_{\|x\| \rightarrow \infty} (x_1 - \bar{x}_2)^2 = \infty$$

$$\lim_{\|x\| \rightarrow \infty} (\bar{x}_1 - x_2)^2 = \infty$$

In cui  $\bar{x}_1$  e  $\bar{x}_2$  sono rispettivamente dei valori fissati di  $x_1$  e  $x_2$ , otteniamo che la funzione va ad infinito separatamente lungo le due componenti. Questo non implica che la funzione sia coerciva.

Osserviamo infatti che si tratta di funzione quadratica con matrice hessiana semidefinita positiva con autovalori pari a  $\lambda_{\min} = 0$  e  $\lambda_{\max} = 1$ .

Utilizzando il Teorema 16, sappiamo che esiste un punto di minimo globale se e solo se esiste una soluzione del sistema  $\nabla q(x) = 0$ , cioè

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

In particolare, in questo caso sono minimi globali tutti i punti sulla retta  $x_1 = x_2$ , cioè del tipo  $(\xi, \xi)^T$ .

**Esempio 9** Sia data la funzione quadratica

$$q(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - x_1x_2 + x_2x_3 - x_1x_3$$

Trovare, se esiste, il minimo globale.

**Soluzione.** L'annullamento del gradiente conduce al sistema lineare omogeneo

$$\begin{cases} 2x_1 - x_2 - x_3 = 0 \\ -x_1 + 2x_2 + x_3 = 0 \\ -x_1 + x_2 + 2x_3 = 0 \end{cases}$$

La matrice dei coefficienti é la matrice hessiana

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

Poichè  $\nabla^2 f$  é non singolare e definita positiva, l'unica soluzione del sistema  $\nabla q = 0$  é l'unico punto di minimo globale  $x = (0, 0, 0)^T$ .

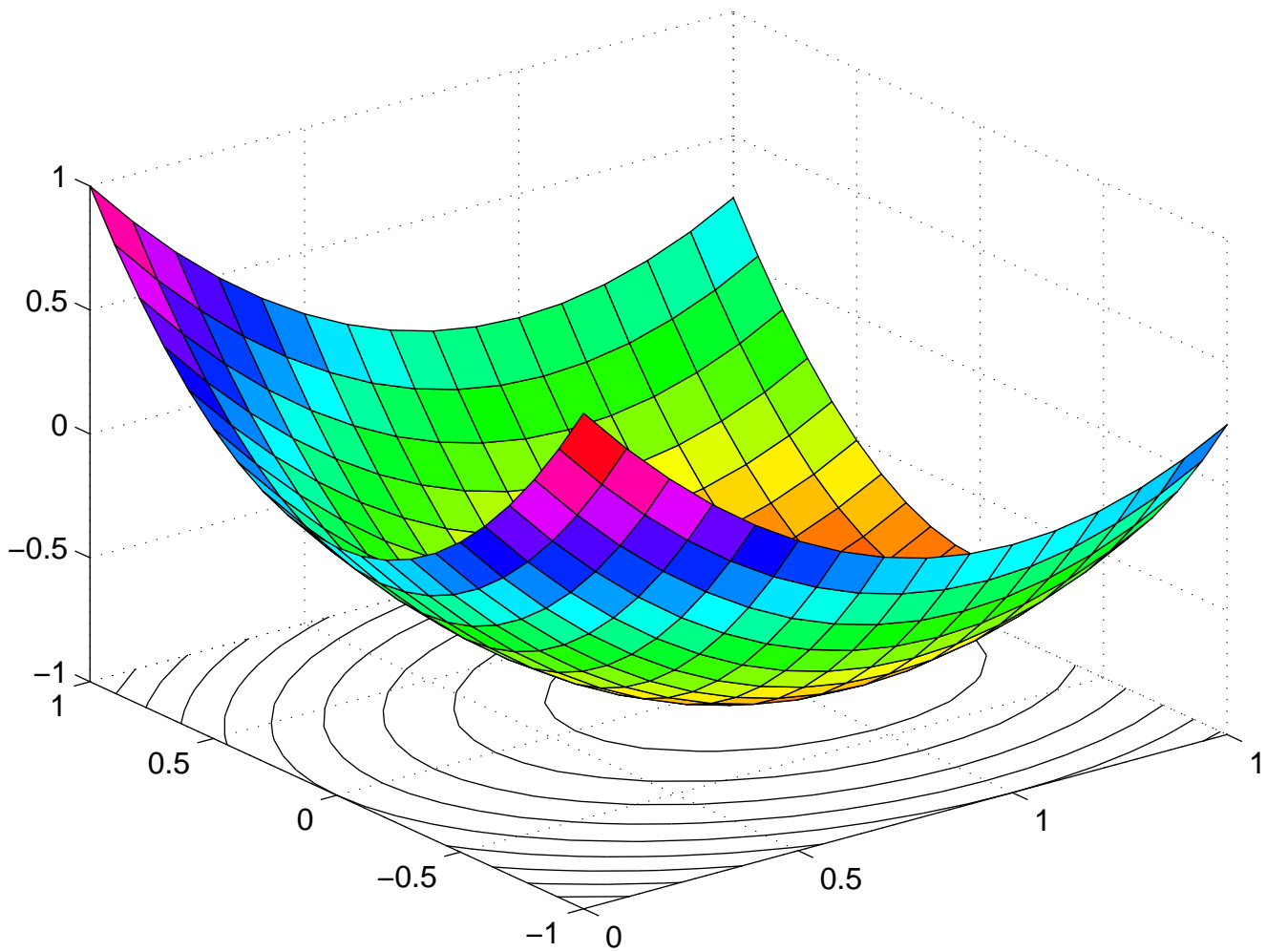


Figure 8: Grafico di  $q(x_1, x_2)$  per  $\alpha = \beta = 1$ .

**Esempio 10** [2] *Sia data la funzione*

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

*Studiare le esistenza e la natura dei punti estremali al variare dei parametri  $\alpha$  e  $\beta$ .*

**Soluzione.**

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

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

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

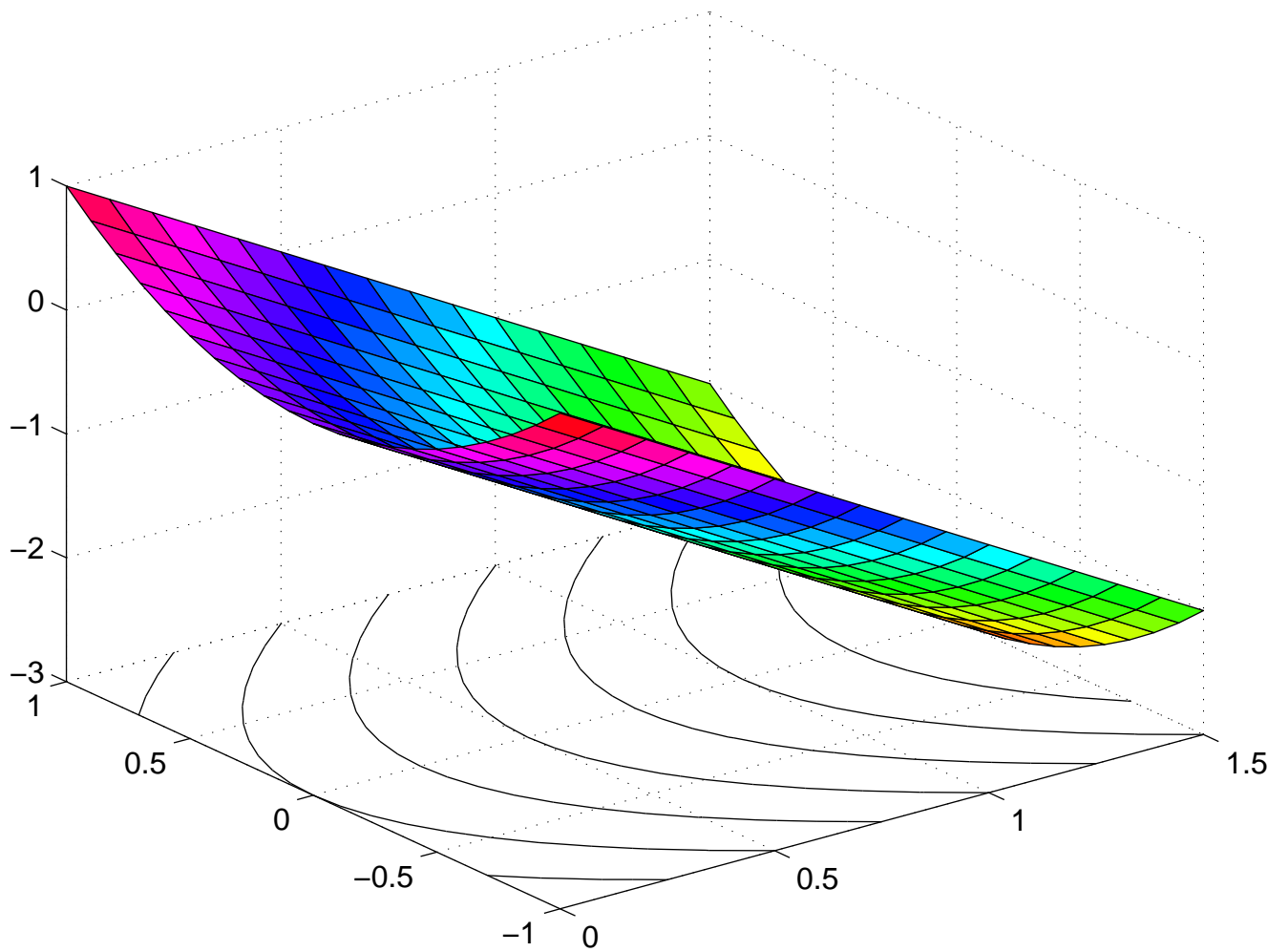


Figure 9: Grafico di  $q(x_1, x_2)$  per  $\alpha = 0$   $\beta = 1$ .

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

Se  $\alpha < 0$  e  $\beta > 0$  si ha un'unica soluzione  $\left(\frac{1}{\alpha}, 0\right)$ . Ma la matrice hessiana é 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 é  $\left(\frac{1}{\alpha}, 0\right)^T$ . Inoltre la matrice  $\nabla^2 q$  é definita negativa; si tratta quindi dell'unico punto di massimo globale.

**Esempio 11** Sia data la funzione

$$q(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2.$$

*Studiare la natura degli eventuali punti estremali.*

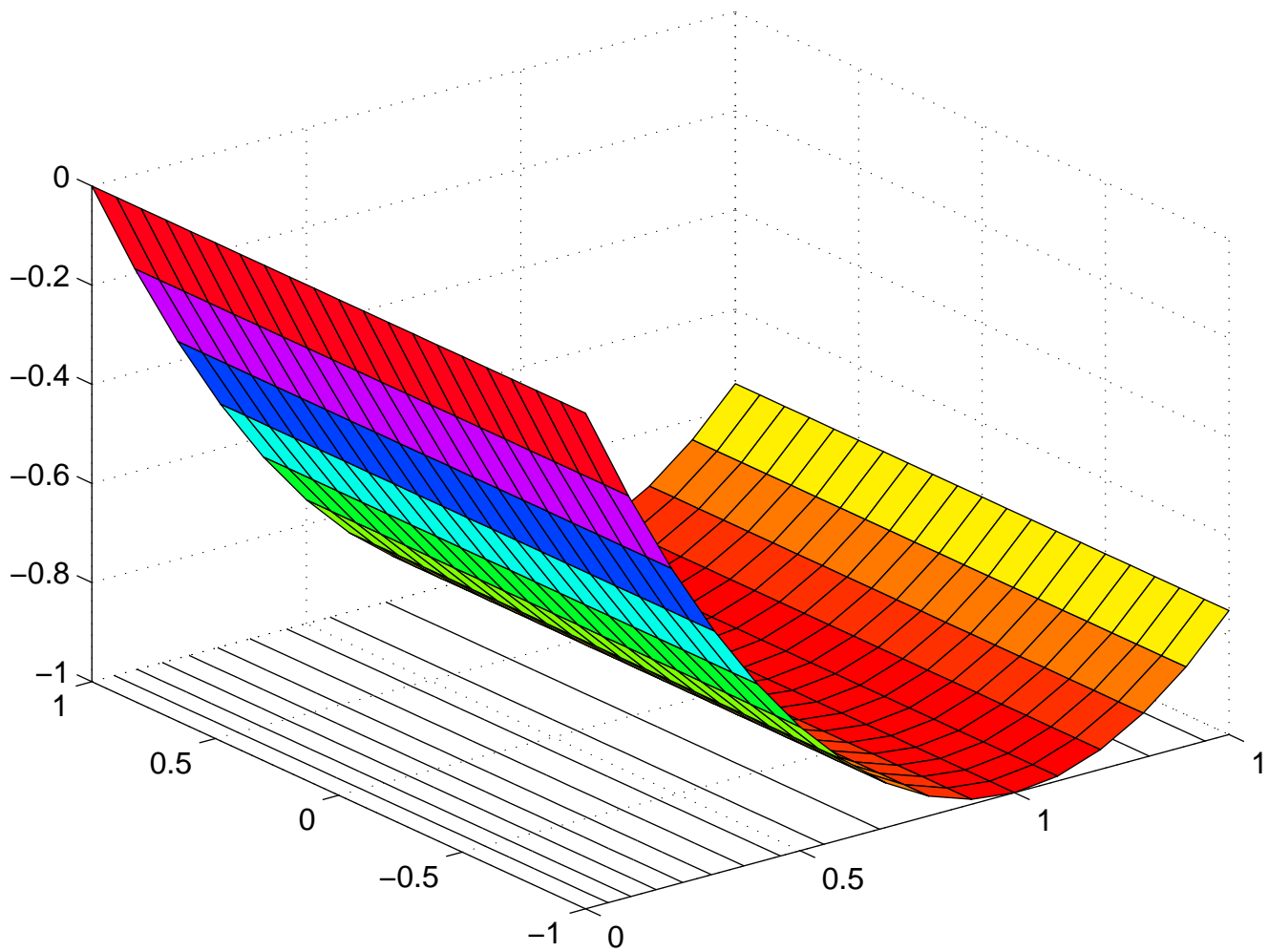


Figure 10: Grafico di  $q(x_1, x_2)$  per  $\alpha = 1$   $\beta = 0$ .

**Soluzione.** Risulta

$$\nabla q(x) = Qx + c = \begin{pmatrix} 4x_1 + 2x_2 + 1 \\ 2x_2 + 2x_1 - 1 \end{pmatrix}$$

con

$$Q = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$$

La matrice  $Q$  è definita positiva. Infatti  $\lambda_{\min} = 3 - \sqrt{5}$  e  $\lambda_{\max} = 3 + \sqrt{5}$ . Esiste quindi un'unica soluzione ottima che si ottiene annullando il gradiente:

$$\begin{aligned} 4x_1 + 2x_2 &= 1 \\ 2x_2 + 2x_1 &= -1 \end{aligned}$$

Si ottiene la soluzione

$$x^* = \begin{bmatrix} -1 \\ 1.5 \end{bmatrix}$$

con valore  $q(x^*) = -5/4$ .

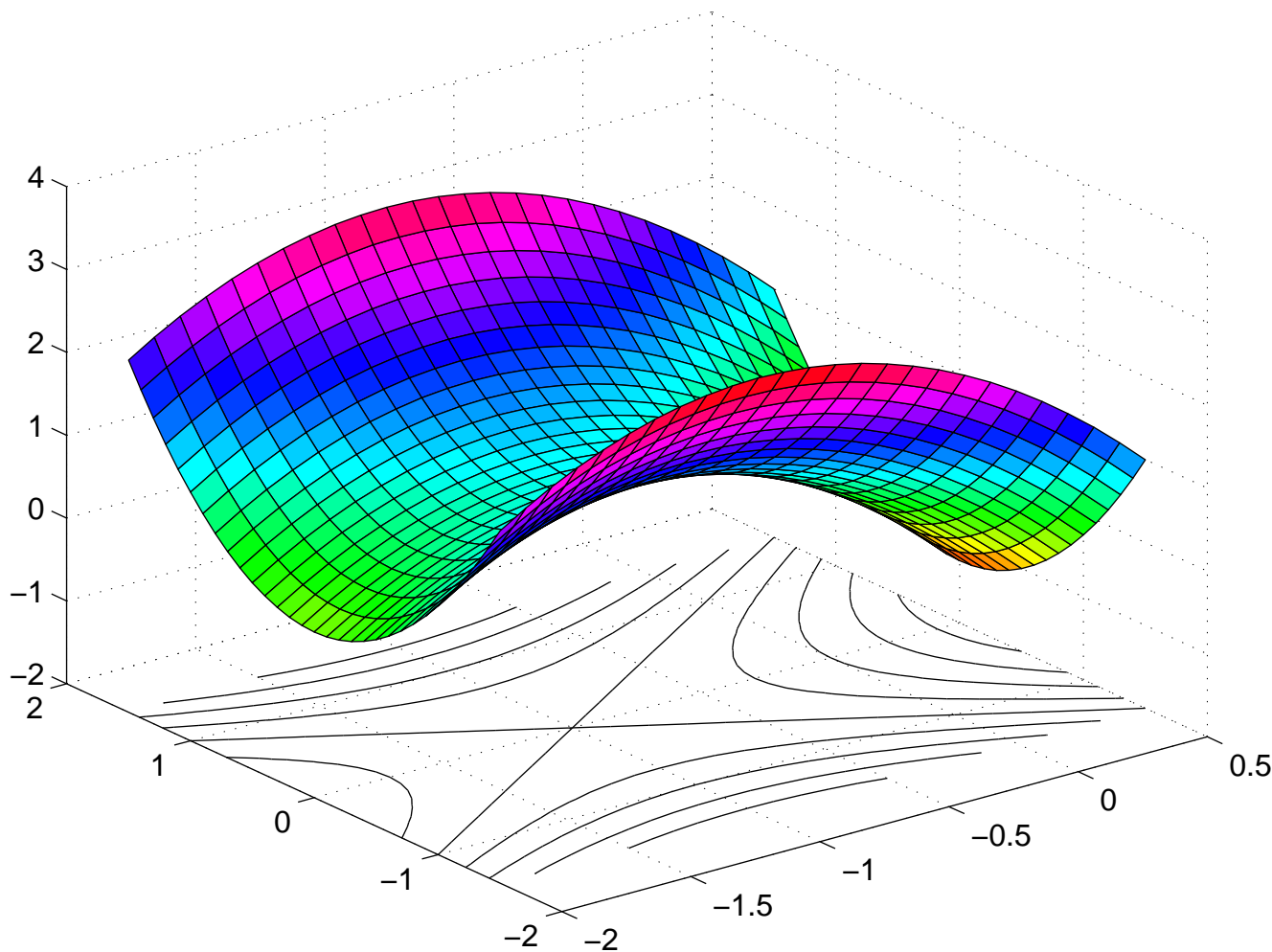


Figure 11: Grafico di  $q(x_1, x_2)$  per  $\alpha = -1$   $\beta = 1$ .

**Esempio 12 (Discriminazione del prezzo [1])** *Un monopolista che produce un unico bene ha due tipi di clienti A e B. Siano  $Q_A$  e  $Q_B$  le quantità offerte dal monopolista ai due clienti. I clienti di tipo A sono disposti a pagare il prezzo  $P_A = f_A(Q_A) = 50 - 5Q_A$  e i clienti di tipo B sono disposti a pagare il prezzo  $P_B = f_B(Q_B) = 100 - 10Q_B$ . Il costo di produzione dipende solo dalla quantità di prodotto finale  $Q = Q_A + Q_B$  ed è  $C = 90 + 20Q$ . Definire il modello di ottimizzazione.*

**Soluzione.** Il profitto (ricavo-costi) è dato dall'espressione:

$$\begin{aligned} f(Q_A, Q_B) &= Q_A P_A + Q_B P_B - [90 + 20(Q_A + Q_B)] \\ &= (50 - 5Q_A)Q_A + (100 - 10Q_B)Q_B - 90 - 20(Q_A + Q_B) \end{aligned}$$

e deve essere massimizzato. Raggruppando e portando in forma standard di minimizzazione, si ottiene il problema

$$\min 5x_1^2 + 10x_2^2 - 30x_1 - 80x_2 - 90$$

avendo indicato con  $x_1 = Q_A$  e  $x_2 = Q_B$ <sup>1</sup>.

<sup>1</sup>Notiamo che sono impliciti i vincoli  $x \geq 0$ .

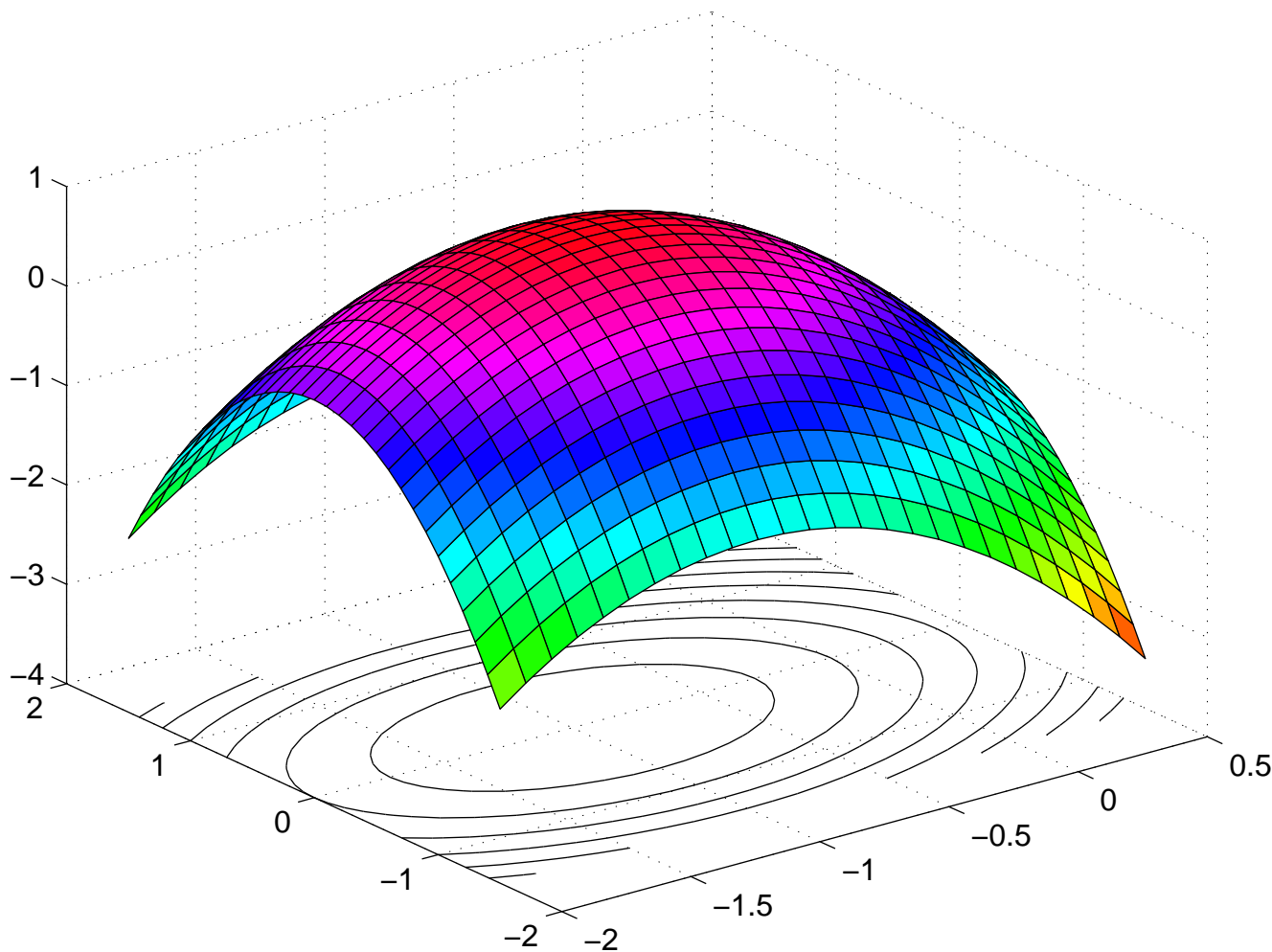


Figure 12: Grafico di  $q(x_1, x_2)$  per  $\alpha = \beta = -1$ .

Si tratta di un problema quadratico con

$$\nabla f = \begin{pmatrix} 10x_1 - 30 \\ 20x_2 - 80 \end{pmatrix} \quad \nabla^2 f = \begin{pmatrix} 10 & 0 \\ 0 & 20 \end{pmatrix}$$

La matrice hessiana è definita positiva, esiste un unico punto di minimo globale che si ottiene dall'annullamento del gradiente.  $\square$

**Esercizio 4** [parte del compito d'esame del 15-9-2003] Dato il problema di programmazione non lineare vincolato

$$\min 3x_1^2x_3 - 27x_1 + x_3^3 + 2x_3x_2^2$$

Dire quali condizioni necessarie soddisfa il punto  $x^* = (3/2, 0, -3)^\top$ .

**Esercizio 5** [parte del compito d'esame del 15-9-2003] Dato il problema di programmazione non lineare non vincolato

$$\min 2x_1^2 + 7x_2^2 - 2x_1x_2 - 4x_1$$

Dire se esiste un punto di minimo (senza determinarlo!).

**Esercizio 6** *Determinare, se esiste, il punto di minimo della funzione*

$$x_1^2 + x_2^2 - 2x_1$$

## 7.5 Il caso vincolato: preliminari

Consideriamo ora il problema vincolato (3), che qui riscriviamo:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & g(x) \leq 0 \\ & h(x) = 0, \end{aligned} \tag{14}$$

con  $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$  e  $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $m \leq n$ . Ricordiamo che

$$\mathcal{F} = \{x \in \mathbb{R}^n : g(x) \leq 0, h(x) = 0\},$$

denota l'insieme ammissibile per il Problema (14).

Diamo alcune definizioni preliminari alla formulazione delle condizioni di ottimalità per il Problema (14).

La prima definizione è quella di *vincoli attivi* e *vincoli non attivi*.

Dato un punto  $x$  ammissibile,  $x \in \mathcal{F}$ , per un qualsiasi vincolo di disuguaglianza  $g_i$  risulterà o  $g_i(x) = 0$  o  $g_i(x) < 0$ . Nel primo caso si dice che il vincolo  $g_i$  è *attivo* in  $x$ ; nel secondo caso si dice che  $g_i$  *non è attivo* in  $x$ . Denotiamo con  $I_a(x)$  l'insieme degli indici dei vincoli di disuguaglianza attivi in  $x$ ; cioè:

$$I_a(x) = \{i \in \{1, \dots, p\} : g_i(x) = 0\}.$$

Sia  $p_a$  la cardinalità dell'insieme  $I_a(x)$ ; è sempre possibile riordinare i vincoli di disuguaglianza in modo tale che risulti  $g_i(x) = 0, i = 1, \dots, p_a$  e  $g_i(x) < 0, i = p_a + 1, \dots, p$ ; e quindi partizionare il vettore  $g$  in due sottovettori, rispettivamente  $g_a$  dei vincoli attivi e  $g_n$  dei vincoli non attivi, il primo di dimensione  $p_a$  e il secondo di dimensione  $p - p_a$ :

$$g(x) = \begin{pmatrix} g_a(x) \\ g_n(x) \end{pmatrix}. \tag{15}$$

Osserviamo che la partizione dipende dal punto  $x \in \mathcal{F}$  considerato: cambiando il punto cambia anche la partizione.

Per quel che riguarda i vincoli di uguaglianza, questi, in un punto ammissibile, sono ovviamente tutti attivi. Possiamo concludere con la seguente considerazione: i vincoli attivi in  $x \in \mathcal{F}$  sono i vincoli di disuguaglianza attivi in  $x$  e tutti i vincoli di uguaglianza.

Una seconda definizione preliminare è quella di *regolarità* dei vincoli in un punto  $x$  ammissibile.

**Definizione 6 (Regolarità dei vincoli)** Diremo che i vincoli sono regolari in un punto  $x \in \mathcal{F}$ , ovvero che  $x \in \mathcal{F}$  è un punto di regolarità per i vincoli, se i gradienti dei vincoli attivi in  $x$  sono tra di loro linearmente indipendenti.

Se la regolarità vale per tutti i punti di  $\mathcal{F}$  diremo che i vincoli sono regolari, senza ulteriori specifiche.

Supponendo di avere riordinato i vincoli di disuguaglianza come nella (15), risulta dalla definizione ora data che i vincoli sono regolari nel punto ammissibile  $x$  se la matrice gradiente dei vincoli attivi:

$$\nabla \begin{bmatrix} g_a(x) \\ h(x) \end{bmatrix} = [\nabla g_a(x) \quad \nabla h(x)], \tag{16}$$

di dimensione  $n \times (p_a + m)$ , ha rango pieno  $p_a + m$ .

Se un punto  $x$  non è di regolarità, i vincoli attivi in  $x$ , *linearizzati*, risultano linearmente dipendenti. Quindi la non regolarità denota una qualche forma di ridondanza dei vincoli nel punto considerato. Se la soluzione del Problema (14) non è un punto di regolarità, questa ridondanza determina una complicazione nella formulazione delle condizioni di ottimalità. Pertanto nel seguito, le condizioni di ottimalità nel caso di vincoli non lineari verranno date con riferimento alle sole soluzioni regolari, che sono quelle che più frequentemente si presentano.

La terza definizione preliminare è quella di *funzione Lagrangiana* associata al Problema (14), o, sinteticamente, *Lagrangiano* del Problema (14).

La funzione Lagrangiana  $L(x, \lambda, \mu)$ , con  $\lambda \in \mathbb{R}^p$  e  $\mu \in \mathbb{R}^m$ , è definita come una combinazione lineare di tutte le funzioni del problema, con coefficienti  $\lambda_i$  per i vincoli di disuguaglianza e  $\mu_j$  per i vincoli di uguaglianza:

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^p \lambda_i g_i(x) + \sum_{j=1}^m \mu_j h_j(x), \quad (17)$$

ovvero, con notazione vettoriale:

$$L(x, \lambda, \mu) = f(x) + \lambda^T g(x) + \mu^T h(x). \quad (18)$$

I coefficienti  $\lambda_i$  e  $\mu_j$  vengono detti, in questo contesto, *moltiplicatori*, rispettivamente dei vincoli di disuguaglianza e di uguaglianza.

## 7.6 Il caso vincolato: vincoli di uguaglianza lineari

Per semplicità cominciamo a considerare il problema di ottimizzazione con soli vincoli di uguaglianza lineari:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & a_j^T x - b_j = 0, \quad j = 1, \dots, m, \end{aligned} \quad (19)$$

con  $a_j \in \mathbb{R}^n$  e  $b_j \in \mathbb{R}$  per ogni  $j$ ; in forma matriciale:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & Ax - b = 0, \end{aligned} \quad (20)$$

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  $b_j$ .

La funzione Lagrangiana per il Problema (19) diviene:

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

o, in forma vettoriale,

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

Allo scopo di definire condizioni di ottimo per problemi vincolati è necessario introdurre la seguente definizione.

**Definizione 7 (Direzione ammissibile)** Un vettore  $d \in \mathbb{R}^n$  è una direzione ammissibile nel punto  $x \in \mathcal{F}$  se esiste uno scalare  $\alpha^{\max} > 0$  tale che risulti:

$$x + \alpha d \in \mathcal{F} \text{ per ogni } \alpha \in (0, \alpha^{\max}]. \quad (21)$$

Quindi, se  $d$  è una direzione ammissibile in un punto, spostandosi da questo punto lungo  $d$  di una quantità sufficientemente piccola, si è sicuri di poter ottenere un nuovo punto ammissibile.

Si può allora dare una semplice condizione necessaria di ottimo vincolato basata sulle definizioni di direzione ammissibile e direzione di discesa (data nel paragrafo 7.2).

**Teorema 17** Condizione necessaria affinché  $x^*$  sia una soluzione locale del Problema (19) è che non esista una direzione  $d \in \mathbb{R}^n$  che sia contemporaneamente ammissibile e di discesa per  $f$  nel punto  $x^*$ .

**Dimostrazione.** La dimostrazione segue dal fatto che, se esiste una direzione  $\bar{d}$  tale che il punto  $x^* + \alpha \bar{d} \in \mathcal{F}$  per  $\alpha \in (0, \alpha_{\max}^{(2)}]$  e  $f(x^* + \alpha \bar{d}) < f(x^*)$  per  $\alpha \in (0, \alpha_{\max}^{(1)}]$ , allora è possibile trovare un valore dello spostamento  $\alpha_{\max} \leq \min\{\alpha_{\max}^{(1)}, \alpha_{\max}^{(2)}\}$  per cui il punto  $x^* + \alpha \bar{d}$  sia contemporaneamente ammissibile e con un valore della funzione  $f(x^* + \alpha \bar{d}) < f(x^*)$  per ogni  $\alpha \in (0, \alpha_{\max}]$  contraddicendo il fatto che  $x^*$  sia un punto di minimo locale vincolato per  $f$ .  $\square$

Ovviamente la condizione espressa dal Teorema 17 è una condizione necessaria di ottimo anche per il problema nella forma più generale (Problema 1). Tuttavia nel caso generale di vincoli di uguaglianza non lineari tale condizione può essere banalmente soddisfatta perché in un punto potrebbe essere vuoto l'insieme delle direzioni ammissibili.

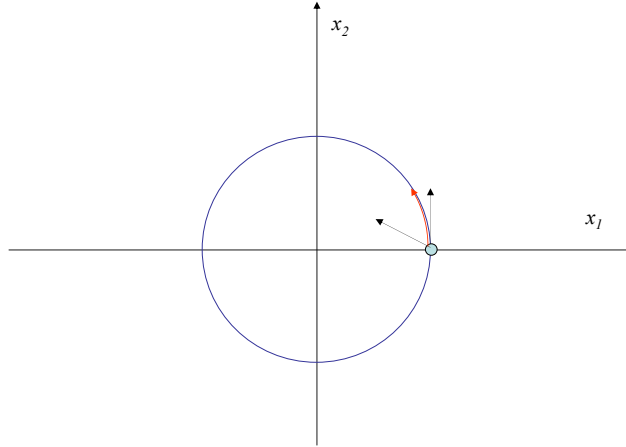


Figure 13: Esempio di vincolo non lineare per cui non esistono direzioni ammissibile  $d$  qualsiasi sia il punto ammissibile.

Ad esempio, nel caso di  $\mathcal{F} = \{x_1^2 + x_2^2 = 1\}$ , in ogni punto ammissibile non esiste una direzione ammissibile  $d$ . Per mantenere l'ammissibilità è necessario spostarsi lungo la curva definita dal vincolo stesso (vedi Figura 13).

Possiamo caratterizzare le direzioni ammissibili per il problema con vincoli di uguaglianza lineari.

**Teorema 18** *Sia  $x$  un punto ammissibile per il Problema (20). Il vettore  $d \in \mathbb{R}^n$  è una direzione ammissibile in  $x$  se e solo se:*

$$Ad = 0. \quad (22)$$

**Dimostrazione.** Consideriamo il punto  $x + \alpha d$  ottenuto muovendosi lungo  $d$  della quantità  $\alpha$ ; poiché  $Ad = 0$ , si ha

$$A(x + \alpha d) = Ax + \alpha Ad = b.$$

□

Ad esempio, consideriamo il sistema di equazioni lineari

$$\begin{aligned} x_1 + x_2 - x_3 &= 1 \\ x_1 - x_2 &= 2 \end{aligned} \quad (23)$$

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 soluzioni del tipo  $d = (t, t, 2t)^T = t(1, 1, 2)^T$  con  $t \in \mathbb{R}$ .

**Osservazione 1** *Notiamo che nel caso lineare una direzione  $d$  che soddisfa la (22) è ammissibile qualunque sia il punto  $x$  ammissibile, e l'ammissibilità si conserva per ogni valore di  $\alpha$ .*

Quindi nel caso di vincoli di uguaglianza lineari è possibile caratterizzare le direzioni ammissibili. Possiamo allora particularizzare la Proposizione 17 ricordando la caratterizzazione delle direzioni di discesa del Teorema 10 e ottenere la seguente condizione necessaria di ottimo.

**Teorema 19** *Condizione necessaria affinché  $x^*$  sia una soluzione locale del Problema (20) è che risulti*

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

Pertanto, mentre nel caso non vincolato la condizione  $\nabla f(x^*)^T d \geq 0$  deve valere per ogni  $d \in \mathbb{R}^n$ , 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 (22).

Per ottenere una condizione di più pratico impiego, conviene approfondire l'analisi del vincolo lineare. A tale scopo, indichiamo con  $A_1, A_2, \dots, A_n$  le colonne della matrice  $A$ , e scriviamo il vincolo mettendole in evidenza:

$$A_1 x_1 + A_2 x_2 + \dots + A_n x_n - b = 0. \quad (24)$$

Osserviamo poi che risulta:

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

Allo scopo di semplificare la trattazione, supponiamo che i vincoli del Problema (20) siano regolari. Si tratta di un'ipotesi di lavoro che semplifica la dimostrazione delle condizioni di ottimo, ma che non è effettivamente necessaria.

La condizione di regolarità dei vincoli in un punto ammissibile, nel caso di vincoli lineari, si traduce nella condizione che la matrice  $A^T$ , e quindi

la matrice  $A$  abbia rango  $m$ .

Notiamo che in questo caso la condizione non dipende dal punto ammissibile considerato, e quindi potremo semplicemente parlare di regolarità dei vincoli.

Possiamo allora estrarre dalla matrice  $A$  un numero  $m$  di colonne linearmente indipendenti, e possiamo riordinare le colonne  $A_1, A_2, \dots, A_n$  e di conseguenza le variabili  $x_1, x_2, \dots, x_n$  in modo tale che quelle linearmente indipendenti siano le prime  $m$ . Possiamo cioè riscrivere il sistema (24), con l'intesa che le colonne  $A_1, A_2, \dots, A_m$  siano linearmente indipendenti.

Con questa intesa, introduciamo le notazioni:

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

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

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 (24) nella forma:

$$Bx_B + Nx_N - b = 0, \quad (25)$$

e riscrivere la (22) nella forma:

$$Bd_B + Nd_N = 0. \quad (26)$$

Concludiamo questa analisi preliminare osservando che, poichè  $B$  è non singolare, dalla (26) si ottiene:

$$d_B = -B^{-1}Nd_N; \quad (27)$$

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

$$d = \begin{bmatrix} -B^{-1}Nd_N \\ d_N \end{bmatrix}, \quad (28)$$

ove  $d_N \in \mathbb{R}^{(n-m)}$  è arbitrario.

Ad esempio, con riferimento al sistema (23), 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

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

Possiamo ora enunciare la seguente condizione necessaria di ottimalità per il Problema (19).

**Teorema 20** *Condizione necessaria affinché  $x^*$  sia una soluzione locale del Problema (19) è che esistano dei moltiplicatori  $\mu_1^*, \mu_2^*, \dots, \mu_m^*$  tali che:*

$$\nabla_x L(x^*, \mu^*) = \nabla f(x^*) + \sum_{j=1}^m \mu_j^* a_j = 0, \quad (29)$$

o, in forma matriciale:

$$\nabla_x L(x^*, \mu^*) = \nabla f(x^*) + A^T \mu^* = 0. \quad (30)$$

**Dimostrazione.** Nella dimostrazione faremo uso dell'ipotesi semplificativa  $\text{rango}(A) = m$ . In questa ipotesi possiamo considerare la partizione di  $A = (B \ N)$  e partizionare di conseguenza anche il gradiente di  $f$  nella forma seguente:

$$\nabla f(x^*) = \begin{bmatrix} \nabla_B f(x^*) \\ \nabla_N f(x^*) \end{bmatrix}, \quad (31)$$

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 (19):

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

può essere riscritta nella forma:

$$\nabla_B f(x^*)^T d_B + \nabla_N f(x^*)^T d_N \geq 0, \quad \text{per ogni } d_B \in \mathbb{R}^m, \text{ e } d_N \in \mathbb{R}^{(n-m)} : Bd_B + Nd_N = 0.$$

Tenendo conto che  $B$  è non singolare possiamo scrivere, in base alla (27),  $d_B = -B^{-1}Nd_N$  e quindi:

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

Mettendo in evidenza  $d_N$  si ottiene:

$$(-\nabla_B f(x^*)^T B^{-1}N + \nabla_N f(x^*)^T)d_N \geq 0, \quad \text{per ogni } d_N \in \mathbb{R}^{n-m}. \quad (32)$$

ovvero

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

Analogamente a come abbiamo ragionato nel caso non vincolato (vedi Teorema 12), affinché la disuguaglianza precedente valga per ogni  $d_N \in \mathbb{R}^{n-m}$ , deve risultare nullo il vettore che premoltiplica  $d_N$ , deve cioè risultare:

$$\nabla_N f(x^*) - N^T(B^{-1})^T \nabla_B f(x^*) = 0; \quad (33)$$

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

Poniamo ora:

$$\mu^* = -(B^{-1})^T \nabla_B f(x^*), \quad (34)$$

e sostituiamo nella (33); otteniamo:

$$\nabla_N f(x^*) + N^T \mu^* = 0. \quad (35)$$

A questo punto, abbiamo fatto vedere che, se  $x^*$  è una soluzione locale del Problema (19), esiste un vettore  $\mu^*$ , che soddisfa le equazioni (34), (35), equazioni che possiamo riscrivere, ricordando che  $(B^{-1})^T = (B^T)^{-1}$ , nella forma:

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

cioè anche in forma matriciale

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

Ma, ricordando che  $A^T = \begin{pmatrix} B^T \\ N^T \end{pmatrix}$ , le due precedenti equazioni altro non sono che la (30) scritta in forma partizionata rispetto alle variabili di base e non di base.  $\square$

Se  $f(x)$  è due volte continuamente differenziabile, valgono anche le seguenti condizioni del secondo ordine, di cui possiamo notare analogie e differenze rispetto al caso non vincolato: di nuovo, anziché fare riferimento a tutte le direzioni  $d \in \mathbb{R}^n$ , occorre fare riferimento alle sole direzioni ammissibili  $d \in \mathbb{R}^n : Ad = 0$ .

**Teorema 21** *Condizione necessaria affinché  $x^*$  sia una soluzione locale del Problema (19) è che esista un moltiplicatore  $\mu^*$  tale che  $\nabla_x L(x^*, \mu^*) = 0$  e inoltre risulti  $d^T \nabla^2 f(x^*) d \geq 0$  per ogni  $d \in \mathbb{R}^n$  tale che  $Ad = 0$ .*

**Teorema 22** *Condizione sufficiente affinché un punto ammissibile  $x^*$  sia una soluzione locale stretta del Problema (19) è che esista un moltiplicatore  $\mu^*$  tale che  $\nabla_x L(x^*, \mu^*) = 0$ , e che risulti  $d^T \nabla^2 f(x^*) d > 0$  per ogni  $d \in \mathbb{R}^n$  tale che  $Ad = 0$ ,  $d \neq 0$ .*

Osserviamo che, essendo il vincolo lineare, la sua derivata seconda è nulla; pertanto è lecito scrivere le due ultime proposizioni sostituendo  $\nabla_x^2 L(x^*, \mu^*)$  al posto di  $\nabla^2 f(x)$ .

**Osservazione 2** *Concludiamo con l'importante osservazione che l'ipotesi di regolarità dei vincoli è stata utilizzata in questo paragrafo solo allo scopo di introdurre e semplificare il caso generale di vincoli non lineari. In realtà, con una diversa trattazione, si può dimostrare per il caso di vincoli lineari la stessa condizione necessaria della Proposizione 20 senza fare uso della regolarità dei vincoli.*

**Esercizio 7** *Sia dato il problema*

$$\begin{aligned} \min \quad & f(x) \\ & x_1 + x_2 - x_3 = 1 \\ & x_1 - x_2 = 2 \end{aligned} \quad (36)$$

1. data  $f(x) = x_1^2 + x_2^2 + x_3^2$ , verificare quali condizioni soddisfa il punto  $x^* = (4/3 - 2/3 - 1/3)^T$  (risposta: minimo locale stretto);

2. data  $f(x) = x_1^2 - x_2^2 + x_3^2$ , verificare quali condizioni soddisfa il punto  $x^* = (1 \ -1 \ -1)^T$  (risposta: minimo locale stretto);
3. data  $f(x) = x_1^2 + x_2^2 - x_3^2$ , verificare quali condizioni soddisfa il punto  $x^* = (2 \ 0 \ 1)^T$  (risposta: non è minimo).
4. data  $f(x) = 3x_1 + 5x_2 - x_3$  verificare che non sono mai verificate le condizioni necessarie del primo ordine.

**Esercizio 8** Sia dato il problema con vincoli di disuguaglianza lineari

$$\min_{x \in \mathbb{R}^n} f(x) \\ a_j^T x - b_j \leq 0, \quad j = 1, \dots, p,$$

ovvero in forma matriciale

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

con  $A$  matrice  $p \times n$  e  $b \in \mathbb{R}^p$ .

Mostrare che, in un punto ammissibile  $\bar{x}$ , i vettori  $d \in \mathbb{R}^n$  tali che

$$a_j^T d \leq 0 \quad \text{per } j \in I_a(\bar{x}) = \{i : a_i^T \bar{x} = b_i\}$$

sono tutte e sole le direzioni ammissibili.

## 7.7 Il caso vincolato: vincoli di uguaglianza non lineari

Consideriamo ora il problema di ottimizzazione con soli vincoli di uguaglianza, in generale non lineari:

$$\min_{x \in \mathbb{R}^n} f(x) \\ h_j(x) = 0, \quad j = 1, \dots, m, \tag{37}$$

o, in forma vettoriale:

$$\min_{x \in \mathbb{R}^n} f(x) \\ h(x) = 0, \tag{38}$$

con  $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ .

La funzione Lagrangiana per il Problema (37) è quindi:

$$L(x, \mu) = f(x) + \sum_{j=1}^m \mu_j h_j(x),$$

o, in forma vettoriale,

$$L(x, \mu) = f(x) + \mu^T h(x).$$

Sia  $x^*$  una soluzione locale del problema, e consideriamo il punto  $x^* + \alpha d$  ottenuto spostandosi da  $x^*$  lungo la direzione  $d \in \mathbb{R}^n$  di una quantità  $\alpha > 0$ ; diremo che lo spostamento  $\alpha d$  è ammissibile se  $h(x^* + \alpha d) = 0$ .

Nel caso di vincoli lineari  $Ax - b = 0$ , abbiamo visto che ogni spostamento  $\alpha d$  tale che  $Ad = 0$  risulta ammissibile. Nel caso di vincoli non lineari, per caratterizzare gli spostamenti ammissibili, dobbiamo approfondire l'analisi, basandosi però in parte su quanto detto per il caso lineare.

A tale proposito, consideriamo la linearizzazione  $h_l(x) = 0$  dei vincoli nell'intorno di  $x^*$ :

$$h_l(x^* + \alpha d) = h(x^*) + \alpha \nabla h(x^*)^T d = 0, \quad (39)$$

e osserviamo che, se  $\nabla h(x^*)^T d = 0$  (che corrisponde a  $Ad = 0$ ) risulta  $h_l(x^* + \alpha d) = 0$ , e quindi  $\alpha d$  è uno spostamento ammissibile nel punto  $x^*$  per il vincolo linearizzato. In generale però non è ammissibile per il vincolo  $h(x)$ , in quanto risulta  $h(x^* + \alpha d) \neq 0$  (vedi la Figura 13 del paragrafo precedente). Tuttavia, nel caso che  $x^*$  sia un punto di regolarità, la linearizzazione del vincolo ci consente di caratterizzare degli spostamenti ammissibili per il vincolo, ricorrendo al *teorema delle funzioni implicite*.

Supponiamo quindi che valga l'ipotesi di regolarità dei vincoli, ovvero

$$\text{rango}(\nabla h(x^*)) = m \leq n.$$

Allora possiamo estrarre dalla matrice gradiente dei vincoli  $\nabla h(x^*)$   $m$  righe linearmente indipendenti, ovvero possiamo scrivere

$$\nabla h(x^*) = \begin{pmatrix} \nabla_B h(x^*) \\ \nabla_N h(x^*) \end{pmatrix} \quad (40)$$

dove  $\nabla_B h(x^*)$  è una sottomatrice di base  $m \times m$  non singolare, costituita dal gradiente di  $h$  rispetto alle variabili di base  $x_B \in \mathbb{R}^m$  e  $\nabla_N h(x^*)$ , una sottomatrice  $(n - m) \times m$  non di base, costituita dal gradiente di  $h$  rispetto alle variabili non di base  $x_N \in \mathbb{R}^{n-m}$ .

Ragionando come abbiamo fatto sull'equazione  $Ad = 0$  nel caso di vincoli lineari, possiamo riscrivere l'equazione

$$\nabla h(x^*)^T d = 0$$

mettendo in evidenza le sottomatrici  $\nabla_B h(x^*)^T$  e  $\nabla_N h(x^*)^T$ , e in corrispondenza le direzioni di base  $d_B \in \mathbb{R}^m$  e non di base  $d_N \in \mathbb{R}^{(n-m)}$ . Ricordando che  $\nabla h(x^*)^T = [\nabla_B h(x^*)^T \quad \nabla_N h(x^*)^T]$ , si ha:

$$\nabla_B h(x^*)^T d_B + \nabla_N h(x^*)^T d_N = 0. \quad (41)$$

La direzione  $d = \begin{pmatrix} d_B \\ d_N \end{pmatrix}$  con  $d_B = -(\nabla_B h(x^*)^T)^{-1} \nabla_N h(x^*)^T d_N$  e  $d_N \in \mathbb{R}^{n-m}$  qualsiasi risulta dunque essere ammissibile per il vincolo linearizzato  $h_l(x)$ , ma non per il vincolo  $h(x)$ .

Il teorema delle funzioni implicite, nel caso che stiamo considerando, ci consente di affermare quanto segue:

**Teorema 23** *Siano  $h : \mathbb{R}^m \times \mathbb{R}^{(n-m)} \rightarrow \mathbb{R}^m$  una funzione delle variabili  $x_B \in \mathbb{R}^m$  e  $x_N \in \mathbb{R}^{(n-m)}$  e  $x^* = (x_B^* \quad x_N^*)^T \in \mathbb{R}^n$  tali che:*

1.  $h(x_B^*, x_N^*) = 0$ ;
2.  $h$  è continuamente differenziabile e risulta  $\nabla_B h(x^*)$  non singolare.

Allora esistono un intorno  $\mathcal{S}_B \subseteq \mathbb{R}^m$  di  $x_B^*$  e  $\mathcal{S}_N \subseteq \mathbb{R}^{(n-m)}$  di  $x_N^*$  e una funzione  $\phi : \mathcal{S}_N \rightarrow \mathcal{S}_B$  continuamente differenziabile, e tale che, per ogni  $\alpha > 0$  sufficientemente piccolo, risulta:

$$h(x_B^* + \phi(\alpha d_N), x_N^* + \alpha d_N) = 0, \quad \text{per ogni } d_N \in \mathbb{R}^{(n-m)}.$$

Inoltre la funzione  $\phi$  è tale che:

$$\begin{aligned}\phi(0) &= 0, \\ \nabla\phi(0) &= -\nabla_N h(x^*)[\nabla_B h(x^*)]^{-1}.\end{aligned}$$

□

Pertanto, quando  $\nabla_B h(x^*)$  è non singolare, l'esistenza di una direzione ammissibile per il vincolo linearizzato, pur non fornendo direttamente una direzione ammissibile, consente di affermare l'esistenza in  $(x_B^*, x_N^*)$  di uno spostamento ammissibile  $(\phi(\alpha d_N), \alpha d_N)$ , per  $\alpha > 0$  sufficientemente piccolo, comunque si prenda  $d_N \in \mathbb{R}^{(n-m)}$ . Notiamo che della funzione  $\phi$  non viene data l'espressione esplicita; ci basta essere sicuri della sua esistenza, e poter scrivere lo sviluppo di Taylor al primo ordine:

$$\begin{aligned}\phi(\alpha d_N) &= \phi(0) + \alpha \nabla\phi(0)^T d_N + o(\alpha) \\ &= -\alpha([\nabla_B h(x^*)]^{-1})^T \nabla_N h(x^*)^T d_N + o(\alpha),\end{aligned}\quad (42)$$

ove  $o(\alpha)$  denota termini infinitesimi di ordine superiore rispetto ad  $\alpha$ .

Grazie al teorema delle funzioni implicite possiamo enunciare la condizione necessaria di ottimo utilizzando la funzione Lagrangiana come abbiamo fatto nel caso di vincoli di uguaglianza lineari. In questo caso l'ipotesi di regolarità dei vincoli è necessaria per poter enunciare il teorema stesso.

**Teorema 24 (Condizioni necessarie di Lagrange)** *Sia  $x^*$  un punto di regolarità per i vincoli del Problema (37). Condizione necessaria affinché  $x^*$  sia una soluzione locale del problema è che esista un vettore di moltiplicatori  $\mu^*$ , con componenti  $\mu_1^*, \mu_2^*, \dots, \mu_m^*$ , tali che:*

$$\nabla_x L(x^*, \mu^*) = \nabla f(x^*) + \sum_{j=1}^m \mu_j^* \nabla h_j(x^*) = 0, \quad (43)$$

o, in forma matriciale:

$$\nabla_x L(x^*, \mu^*) = \nabla f(x^*) + \nabla h(x^*) \mu^* = 0. \quad (44)$$

**Dimostrazione.** Facciamo vedere che, nelle ipotesi poste, la condizione del teorema corrisponde alla non esistenza di uno spostamento che sia contemporaneamente ammissibile e di discesa. A questo scopo, consideriamo lo spostamento, fornito dal teorema delle funzioni implicite,

$$s = \begin{pmatrix} \phi(\alpha d_N) \\ \alpha d_N \end{pmatrix} \quad \text{per ogni } d_N \in \mathbb{R}^{(n-m)}.$$

Partizioniamo le variabili e i vettori secondo la (40). Quindi

$$x = x^* + s = \begin{bmatrix} x_B^* \\ x_N^* \end{bmatrix} + \begin{bmatrix} \phi(\alpha d_N) \\ \alpha d_N \end{bmatrix} \quad \nabla f(x^*) = \begin{bmatrix} \nabla_B f(x^*) \\ \nabla_N f(x^*) \end{bmatrix}.$$

Possiamo scrivere uno sviluppo di Taylor del primo ordine della funzione  $f$  in  $x = x^* + s$ :

$$f(x^* + s) = f(x_B^* + \phi(\alpha d_N), x_N^* + \alpha d_N) = f(x_B^*, x_N^*) + \nabla_B f(x^*)^T \phi(\alpha d_N) + \nabla_N f(x^*)^T (\alpha d_N) + o(\alpha),$$

e, cioè, tenendo conto della espressione di  $\phi(\alpha d_N)$  data dalla (42):

$$\begin{aligned}f(x^* + s) &= f(x^*) + \nabla_B f(x^*)^T [-\alpha([\nabla_B h(x^*)]^{-1})^T \nabla_N h(x^*)^T d_N + o(\alpha)] + \alpha \nabla_N f(x^*)^T d_N + o(\alpha) = \\ &= f(x^*) + \alpha [-\nabla_B f(x^*)^T ([\nabla_B h(x^*)]^{-1})^T \nabla_N h(x^*)^T + \nabla_N f(x^*)^T] d_N + o(\alpha).\end{aligned}$$

Dalla precedente eguaglianza ricaviamo che, se  $x^*$  è una soluzione locale del Problema (37), dovrà risultare, per ogni  $d_N \in \mathbb{R}^{(n-m)}$ ,

$$[-\nabla_B f(x^*)^T ([\nabla_B h(x^*)]^{-1})^T \nabla_N h(x^*)^T + \nabla_N f(x^*)^T] d_N \geq 0; \quad (45)$$

altrimenti esiste un  $\bar{d}_N$  tale che, per ogni  $\alpha > 0$  sufficientemente piccolo, il punto  $(x_B^* + \phi(\alpha \bar{d}_N), x_N^* + \alpha \bar{d}_N)$  risulta ammissibile, e in questo punto risulta  $f(x_B^* + \phi(\alpha \bar{d}_N), x_N^* + \alpha \bar{d}_N) < f(x_B^*, x_N^*)$ .

(Osservazione: Se si pone  $\nabla_B h(x^*)^T = B, \nabla_N h(x^*)^T = N$ , vediamo che la (45) e la (32) coincidono.)

Si può continuare a ragionare come nella dimostrazione della Proposizione 20. La condizione (45) si può scrivere, trasponendo i vettori,

$$[\nabla_N f(x^*) - \nabla_N h(x^*) \nabla_B h(x^*)^{-1} \nabla_B f(x^*)]^T d_N \geq 0 \quad \text{per ogni } d_N \in \mathbb{R}^{(n-m)}.$$

Affinché questa disequazione sia vera per ogni  $d_N \in \mathbb{R}^{(n-m)}$ , deve risultare

$$\nabla_N f(x^*) - \nabla_N h(x^*) \nabla_B h(x^*)^{-1} \nabla_B f(x^*) = 0.$$

Posto

$$\mu^* = -\nabla_B h(x^*)^{-1} \nabla_B f(x^*),$$

possiamo scrivere le due condizioni

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

che sono esattamente la (44) in forma partizionata □

La condizione necessaria di ottimalità enunciata nella proposizione precedente è nota come *condizione di Lagrange*; i moltiplicatori  $\mu^*$  vengono spesso chiamati *moltiplicatori di Lagrange*.

La condizione necessaria  $\nabla_x L(x^*, \mu^*) = 0$ , e quella di ammissibilità  $h(x^*) = 0$  danno luogo ad un sistema di  $n + m$  equazioni nelle  $n + m$  incognite  $x, \mu$ :

$$\begin{aligned} \nabla_x L(x, \mu) &= 0, \\ h(x) &= 0 \end{aligned}$$

che ha, tra le sue soluzioni, tutte le soluzioni locali regolari del problema (37) e i corrispondenti moltiplicatori. Quindi una via per risolvere il problema (37) è quella di risolvere rispetto a  $x, \mu$  il precedente sistema, ricordando comunque che si ottengono solo le soluzioni regolari. Ovviamente, nei confronti di questo approccio valgono le obiezioni più volte fatte presenti, relative alla difficoltà di risolvere un sistema non lineare di  $n + m$  equazioni in  $n + m$  incognite, ove la dimensione  $n + m$  può essere elevata.

Se  $f(x)$  e  $h(x)$  sono due volte continuamente differenziabile, valgono anche la seguenti condizioni del secondo ordine, cui si riconducono anche quelle del caso lineare, se si nota che per il caso lineare risulta  $\nabla h(x^*)^T = A, \nabla^2 h(x^*) = 0$ .

**Teorema 25 (Condizioni necessarie del 2° ordine)** *Sia  $x^*$  un punto di regolarità per i vincoli del Problema (37). Condizione necessaria affinché  $x^*$  sia una soluzione locale del Problema (37) è che esista un moltiplicatore  $\mu^*$  tale che  $\nabla_x L(x^*, \mu^*) = 0$  e inoltre risulti*

$$d^T \nabla_x^2 L(x^*, \mu^*) d \geq 0 \quad \text{per ogni } d \in \mathbb{R}^n : \nabla h(x^*)^T d = 0.$$

**Teorema 26 (Condizioni sufficienti del 2° ordine)** *Condizione sufficiente affinché un punto ammissibile  $x^*$  sia una soluzione locale stretta del Problema (37) è che esista un moltiplicatore  $\mu^*$  tale che  $\nabla_x L(x^*, \mu^*) = 0$ , e che risulti  $d^T \nabla_x^2 L(x^*, \mu^*) d > 0$  per ogni  $d \in \mathbb{R}^n: \nabla h(x^*)^T d = 0, d \neq 0$ .*

Notiamo che nella condizione sufficiente non è richiesto che  $x^*$  sia un punto di regolarità.

**Esercizio 9** *Sia dato il problema*

$$\begin{aligned} \min \quad & x_1 \\ & (x_1 - 4)^2 + x_2^2 = 16 \end{aligned}$$

1. *Dire quali condizioni di ottimo soddisfa il punto  $(0, 0)^T$  (risposta: è un minimo locale stretto).*
2. *Dire quali condizioni di ottimo soddisfa il punto  $(8, 0)^T$  (risposta: non è un minimo locale).*
3. *Si aggiunga al problema il vincolo  $x_1 = 0$ ; verificare la regolarità dei vincoli nel punto  $(0, 0)^T$  e se si possono applicare le condizioni sufficienti del 2° ordine.*

**Esercizio 10** *Sia dato il problema*

$$\begin{aligned} \min \quad & f(x) \\ & (x_1 - 1)^3 + x_2 = 0 \\ & x_2 = 0. \end{aligned}$$

1. *verificare la regolarità dei vincoli nel punto  $(1, 0)^T$ .*
2. *Sia  $f(x) = -x_1$ , applicare le condizioni di Lagrange e verificare se è possibile applicare le condizioni sufficienti del 2° ordine.*
3. *Sia  $f(x) = -x_2$ , applicare le condizioni di Lagrange e verificare se è possibile applicare le condizioni sufficienti del 2° ordine.*

## 7.8 Analisi di sensibilità per vincoli di uguaglianza

Una volta determinata una soluzione  $x^*$  del Problema (37), risulta di notevole interesse avere indicazioni su come varia il valore ottimo della funzione obiettivo  $f(x^*)$ , se si effettuano (piccole) variazioni dei dati del problema. L'analisi di queste variazioni viene detta *analisi di sensibilità*. Di particolare interesse risulta l'analisi di sensibilità rispetto al secondo membro di un vincolo di uguaglianza  $h_j(x) = 0$ , in quanto mette in evidenza una fondamentale proprietà del moltiplicatore  $\mu_j^*$ .

Sia  $x^*$  una soluzione del problema (37),  $f(x^*)$  il valore ottimo della funzione obiettivo. Vogliamo analizzare cosa accade al valore ottimo della funzione obiettivo se il secondo membro del vincolo varia da 0 a un valore  $\epsilon$ . L'analisi che si vuole effettuare si riferisce quindi alla perturbazione del Problema (37) del tipo:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & h_j(x) = \epsilon_j \quad j = 1, \dots, m \end{aligned}$$

ovvero in forma vettoriale

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & h(x) = \epsilon \end{aligned} \tag{46}$$

con  $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$  ed  $\epsilon \in \mathbb{R}^m$ . Vogliamo cioè analizzare la relazione tra la soluzione trovata per il Problema (37) e quella del problema perturbato, la cui funzione Lagrangiana è data, mettendone in evidenza la dipendenza da  $\epsilon$ , da:

$$L(x, \mu; \epsilon) = f(x) + \mu^T(h(x) - \epsilon).$$

A titolo di esempio consideriamo il problema con un solo vincolo lineare

$$\begin{aligned} \min \quad & x_1^2 + x_2^2 \\ & x_1 + x_2 = 1 \end{aligned}$$

la cui soluzione ottima è  $x^* = (1/2 \ 1/2)^T$  di valore  $f(x^*) = 1/2$ , con moltiplicatore  $\mu^* = -1$ , ottenuta risolvendo il sistema delle condizioni necessarie di Lagrange

$$\begin{aligned} 2x_1^* + \mu^* &= 0 \\ 2x_2^* + \mu^* &= 0 \\ x_1^* + x_2^* &= 1 \end{aligned}$$

Supponiamo di perturbare il termine noto di  $\epsilon$  e consideriamo il nuovo vincolo  $x_1 + x_2 = 1 + \epsilon$  con  $\epsilon \in \mathbb{R}$ . Se un punto  $x^*(\epsilon)$  è ottimo per il problema perturbato (PP) allora esiste un  $\mu^*(\epsilon)$  tale che

$$\begin{aligned} 2x_1^*(\epsilon) + \mu^*(\epsilon) &= 0 \\ 2x_2^*(\epsilon) + \mu^*(\epsilon) &= 0 \\ x_1^*(\epsilon) + x_2^*(\epsilon) &= 1 + \epsilon \end{aligned}$$

Risolvendo si ottiene la soluzione

$$x_1^*(\epsilon) = \frac{1}{2} + \frac{\epsilon}{2}, \quad x_2^*(\epsilon) = \frac{1}{2} + \frac{\epsilon}{2}, \quad \mu^*(\epsilon) = -1 - \epsilon.$$

La funzione obiettivo in questo punto vale

$$f(x^*(\epsilon)) = \frac{1}{2} + \epsilon + \frac{\epsilon^2}{2} = f(x^*) - \mu^* \epsilon + o(\epsilon).$$

Ovvero la differenza  $f(x^*(\epsilon)) - f(x^*)$  ha un andamento che, in prima approssimazione, è del tipo  $-\mu^* \epsilon$ . Questa considerazione può essere generalizzata.

Indichiamo con  $x^*(\epsilon)$  una soluzione del problema (46) e con  $\mu^*(\epsilon)$  il corrispondente moltiplicatore. Per  $\epsilon = 0$  abbiamo  $x^*(0) = x^*$  e  $\mu^*(0) = \mu^*$ . Vogliamo studiare l'andamento della funzione  $f(x^*(\epsilon))$  per piccole variazioni di  $\epsilon$  nell'intorno di  $\epsilon = 0$ . A tale scopo possiamo premettere il seguente risultato, che si riporta senza dimostrazione.

**Teorema 27** *Sia dato il problema (37) e siano  $x^*, \mu^*$  rispettivamente una soluzione e il corrispondente moltiplicatore. Supponiamo che  $x^*$  sia un punto regolare e che  $(x^*, \mu^*)$  soddisfino le condizioni sufficienti del 2° ordine. Si consideri una famiglia di problemi (46), ovvero*

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & h(x) = \epsilon, \end{aligned}$$

parametrizzata dal vettore  $\epsilon \in \mathbb{R}^m$ .

Allora esiste un intorno  $\mathcal{S}$  di  $\epsilon = 0$  tale che per ogni  $\epsilon \in \mathcal{S}$

1. esistono un  $x^*(\epsilon) \in \mathbb{R}^n, \mu^*(\epsilon) \in \mathbb{R}^m$  che sono soluzione e corrispondente moltiplicatore del problema (46);
2. risulta  $x^*(0) = x^*, \mu^*(0) = \mu^*$ ;
3.  $x^*(\epsilon) \in \mathbb{R}^n, \mu^*(\epsilon) \in \mathbb{R}^m$  sono funzioni di  $\epsilon$  continuamente differenziabili in  $\mathcal{S}$ .

Nel seguito, per semplificare l'analisi facciamo dapprima riferimento ad un problema con un solo vincolo di uguaglianza:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & h(x) = 0 \end{aligned} \tag{47}$$

con  $h : \mathbb{R}^n \rightarrow \mathbb{R}$ , la cui soluzione è  $x^*$  con moltiplicatore  $\mu^* \in \mathbb{R}$ .

Il problema perturbato è:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & h(x) = \epsilon, \end{aligned} \tag{48}$$

con  $\epsilon \in \mathbb{R}$ .

Indichiamo con  $x^*(\epsilon)$  una soluzione del Problema (48) e  $\mu^*(\epsilon) \in \mathbb{R}$  il corrispondente moltiplicatore, che, nelle ipotesi del Teorema 27 esistono e sono funzioni di  $\epsilon$  continuamente differenziabili in un intorno di  $\epsilon = 0$ . Inoltre per  $\epsilon = 0$  abbiamo  $x^*(0) = x^*$  e  $\mu^*(0) = \mu^*$ .

Possiamo ora ottenere indicazioni su come varia  $f(x^*(\epsilon))$  per piccole variazioni di  $\epsilon$ , nell'intorno di  $\epsilon = 0$ . Infatti, poiché le funzioni  $f(x^*(\epsilon))$  e  $x^*(\epsilon)$  sono continuamente differenziabili rispetto ad  $\epsilon$  (almeno in un intorno  $\mathcal{S}$  di  $\epsilon = 0$ ), possiamo scrivere la formula di Taylor troncata ai termini del primo ordine:

$$f(x^*(\epsilon)) - f(x^*(0)) = \left. \frac{df(x^*(\epsilon))}{d\epsilon} \right|_{(\epsilon=0)} \epsilon + o(\epsilon).$$

Quindi per avere informazioni sull'andamento  $f(x^*(\epsilon)) - f(x^*)$  si deve determinare il termine  $\left. \frac{df(x^*(\epsilon))}{d\epsilon} \right|_{(\epsilon=0)}$ .

Tale determinazione non è immediata in quanto la relazione tra  $x^*(\epsilon)$  e  $f(x^*(\epsilon))$  si ottiene solo risolvendo il problema (46). Per superare questa difficoltà, possiamo utilizzare la funzione Lagrangiana per il problema (46) che è data da:

$$L(x, \mu, \epsilon) = f(x) + \mu(h(x) - \epsilon).$$

Calcolata in  $(x^*(\epsilon), \mu^*(\epsilon))$ , si ottiene

$$L(x^*(\epsilon), \mu^*(\epsilon), \epsilon) = f(x^*(\epsilon)) + \mu^*(\epsilon)(h(x^*(\epsilon)) - \epsilon) = f(x^*(\epsilon)),$$

in quanto in  $x^*(\epsilon)$  il vincolo del problema perturbato deve essere soddisfatto. Da cui si ottiene che

$$\left. \frac{dL(x^*(\epsilon), \mu^*(\epsilon), \epsilon)}{d\epsilon} \right|_{(\epsilon=0)} = \left. \frac{df(x^*(\epsilon))}{d\epsilon} \right|_{(\epsilon=0)}$$

Si tratta quindi di determinare l'espressione di  $\left. \frac{dL(x^*(\epsilon), \mu^*(\epsilon), \epsilon)}{d\epsilon} \right|_{(\epsilon=0)}$ . Si tratta di una funzione composta, e quindi si può scrivere

$$\begin{aligned} \left. \frac{dL(x^*(\epsilon), \mu^*(\epsilon), \epsilon)}{d\epsilon} \right|_{(\epsilon=0)} &= \left[ \nabla_x L(x^*(\epsilon), \mu^*(\epsilon), \epsilon)^T \frac{dx^*(\epsilon)}{d\epsilon} \right]_{(\epsilon=0)} \\ &+ \left[ \frac{\partial L(x^*(\epsilon), \mu^*(\epsilon), \epsilon)}{\partial \mu} \frac{d\mu^*(\epsilon)}{d\epsilon} \right]_{(\epsilon=0)} + \left. \frac{\partial L(x^*(\epsilon), \mu^*(\epsilon), \epsilon)}{\partial \epsilon} \right|_{(\epsilon=0)} \end{aligned}$$

Osserviamo che

$$\left[ \nabla_x L(x^*(\epsilon), \mu^*(\epsilon), \epsilon) \right]_{(\epsilon=0)} = \nabla_x L(x^*(0), \mu^*(0), 0) = \nabla_x L(x^*, \mu^*) = 0$$

perché  $x^*, \mu^*$  soddisfano le condizioni di Lagrange; inoltre

$$\left. \frac{\partial L(x^*(\epsilon), \mu^*(\epsilon), \epsilon)}{\partial \mu} \right|_{(\epsilon=0)} = h(x^*(0)) = h(x^*) = 0$$

in quanto  $x^*$  è ammissibile. Inoltre dall'espressione della funzione Lagrangiana per il problema perturbato, possiamo scrivere

$$\frac{\partial L(x^*(\epsilon), \mu^*(\epsilon), \epsilon)}{\partial \epsilon} = -\mu^*(\epsilon).$$

Otteniamo quindi

$$\left. \frac{df(x^*(\epsilon))}{d\epsilon} \right|_{\epsilon=0} = \left. \frac{dL(x^*(\epsilon), \mu^*(\epsilon), \epsilon)}{d\epsilon} \right|_{(\epsilon=0)} = -\mu^*(0) = -\mu^*. \quad (49)$$

Possiamo allora scrivere:

$$f(x^*(\epsilon)) = f(x^*) - \mu^* \epsilon + o(\epsilon),$$

e cioè, per piccoli valori di  $\epsilon$ , possiamo dire

$$\frac{f(x^*(\epsilon)) - f(x^*)}{\epsilon} \simeq -\mu^*. \quad (50)$$

Questo risultato si esprime dicendo che il moltiplicatore  $\mu^*$ , cambiato di segno, fornisce il *coefficiente di sensibilità* del valore ottimo della funzione obiettivo rispetto a variazioni del termine noto del vincolo.

Questa interpretazione del moltiplicatore  $\mu^*$  è di grande utilità pratica. Infatti se risulta  $\mu^* > 0$  ci si può aspettare che, per piccoli valori di  $\epsilon > 0$  risulti  $f(x^*(\epsilon)) < f(x^*)$ . Si può allora risolvere il Problema (48) con un valore  $\epsilon$  assegnato, per verificare se il *beneficio*, espresso dalla differenza  $f(x^*) - f(x^*(\epsilon))$  compensa il *costo* da sostenere per variare da 0 ad  $\epsilon$  il secondo membro del vincolo <sup>2</sup>.

Nel caso in cui i vincoli siano  $m$  e siano  $\mu_j^*$  il moltiplicatore associato al vincolo  $h_j$ , e  $\epsilon_j$  la variazione del termine noto dello stesso vincolo, possiamo scrivere lo sviluppo di Taylor al primo ordine:

$$f(x^*(\epsilon)) = f(x^*(0)) + \nabla_{\epsilon}^T f(x^*(\epsilon)) |_{(\epsilon=0)} \epsilon + o(\epsilon),$$

dove  $\nabla_{\epsilon} f(x^*(\epsilon))$  è il vettore di  $\mathbb{R}^m$  che denota le derivate parziali prime rispetto a  $\epsilon_j$  con  $j = 1 \dots, m$ . Otteniamo quindi la relazione:

$$f(x^*(\epsilon)) - f(x^*) \simeq - \sum_{j=1}^m \mu_j^* \epsilon_j.$$

---

<sup>2</sup>Ad esempio, in un problema di allocazione di risorse, il vincolo esprime una limitazione sulla disponibilità di risorse; effettuando l'analisi di sensibilità si può valutare se è conveniente aumentare o diminuire la disponibilità.

## 7.9 Il caso vincolato: vincoli di disuguaglianza

Il problema con vincoli di disuguaglianza:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & g_i(x) \leq 0, \quad i = 1, \dots, p, \end{aligned} \quad (51)$$

o, in forma vettoriale:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & g(x) \leq 0, \end{aligned} \quad (52)$$

può essere trattato basandosi largamente su quanto detto nei due paragrafi precedenti.

La funzione Lagrangiana per il Problema (51) è data da:

$$L(x, \lambda) = f(x) + \sum_{i=1}^p \lambda_i g_i(x),$$

o, in forma vettoriale,

$$L(x, \lambda) = f(x) + \lambda^T g(x).$$

Sia  $x^*$  una soluzione locale del problema, e sia  $g_i$  un vincolo non attivo in  $x^*$ , cosicchè risulti  $g_i(x^*) < 0$ . Consideriamo il punto  $x^* + \alpha d$  ottenuto spostandosi da  $x^*$  lungo una qualsiasi direzione  $d \in \mathbb{R}^n$  di una quantità  $\alpha > 0$ ; per  $\alpha$  sufficientemente piccolo risulterà ancora  $g_i(x^* + \alpha d) < 0$ , e quindi per ogni vincolo non attivo, nel punto  $x^*$  qualunque direzione  $d \in \mathbb{R}^n$  risulta ammissibile. In particolare, se tutti i vincoli fossero non attivi in  $x^*$ , a partire da  $x^*$  ci si potrebbe spostare in qualunque direzione senza violare nessun vincolo, e quindi le condizioni necessarie di ottimalità per  $x^*$  sarebbero le stesse del caso non vincolato.

Possiamo concludere che se  $x^*$  è una soluzione locale del Problema (51), è una soluzione locale anche del problema che si ottiene dal Problema (51) eliminando tutti i vincoli non attivi in  $x^*$ , e cioè del problema:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & g_i(x) \leq 0, \quad i \in I_a(x^*), \end{aligned} \quad (53)$$

ove  $I_a(x^*) = \{i \in \{1, \dots, p\} : g_i(x^*) = 0\}$  denota l'insieme degli indici dei vincoli attivi in  $x^*$ .

D'altra parte, in una soluzione locale, i vincoli di disuguaglianza attivi possono essere trattati come vincoli di uguaglianza; pertanto se  $x^*$  è una soluzione locale del Problema (53), è una soluzione locale anche per il Problema:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & g_i(x) = 0, \quad i \in I_a(x^*), \end{aligned} \quad (54)$$

che è un problema con soli vincoli di uguaglianza. Perciò, ricordando la Proposizione 24, possiamo affermare che, se  $x^*$  è un punto di regolarità, esistono moltiplicatori  $\lambda_i, i \in I_a(x^*)$ , tali che:

$$\nabla f(x^*) + \sum_{i \in I_a(x^*)} \lambda_i^* \nabla g_i(x^*) = 0. \quad (55)$$

Assumiamo ora:

$$\lambda_i^* = 0 \text{ per ogni } i : g_i(x^*) < 0,$$

e cioè assegniamo un valore nullo ai moltiplicatori associati ai vincoli non attivi in  $x^*$ ; questa condizione può essere espressa ponendo  $\lambda_i^* g_i(x^*) = 0$ ,  $i \notin I_a(x^*)$ . Osserviamo poi che anche per i vincoli attivi risulta  $\lambda_i^* g_i(x^*) = 0$ ; possiamo allora riscrivere la (55) nella forma:

$$\begin{aligned} \nabla f(x^*) &+ \sum_{i \in I_a(x^*)} \lambda_i^* \nabla g_i(x^*) + \sum_{i \notin I_a(x^*)} \lambda_i^* \nabla g_i(x^*) \\ &= \nabla f(x^*) + \sum_{i=1}^p \lambda_i^* \nabla g_i(x^*) = 0, \end{aligned} \quad (56)$$

con la condizione che

$$\lambda_i^* g_i(x^*) = 0, \quad i = 1, \dots, p.$$

Una importante proprietà dei moltiplicatori  $\lambda_i^*$  associati ai vincoli di disuguaglianza, proprietà non ancora messa in evidenza, è che devono essere tutti non negativi; deve cioè risultare:

$$\lambda_i^* \geq 0, \quad i = 1, \dots, p. \quad (57)$$

Giustificiamo questa proprietà utilizzando l'analisi di sensibilità effettuata in precedenza. Supponiamo di avere nel Problema (51) il solo vincolo  $g(x) \leq 0$ , con  $g : \mathbb{R}^n \rightarrow \mathbb{R}$ :

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & g(x) \leq 0, \end{aligned} \quad (58)$$

sia  $x^*$  una soluzione locale, e sia  $\lambda^*$  il moltiplicatore che soddisfa, con  $x^*$ , le condizioni:

$$\begin{aligned} \nabla f(x^*) + \lambda^* \nabla g(x^*) &= 0, \\ \lambda^* g(x^*) &= 0. \end{aligned}$$

Se in  $x^*$  il vincolo non è attivo, già abbiamo  $\lambda^* = 0$ . Se il vincolo è attivo, per quanto detto in precedenza,  $x^*$ ,  $\lambda^*$  sono anche soluzione e moltiplicatore per il problema:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & g(x) = 0. \end{aligned} \quad (59)$$

Se modifichiamo il Problema (58) introducendo una variazione  $\epsilon > 0$  nel secondo membro del vincolo, otteniamo il problema:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & g(x) \leq \epsilon, \end{aligned} \quad (60)$$

il cui insieme ammissibile  $\mathcal{F}(\epsilon) = \{x : g(x) \leq \epsilon\}$  per  $\epsilon > 0$ , contiene l'insieme ammissibile del Problema (59)  $\mathcal{F} = \{x : g(x) = 0\}$ . Se  $x^*(\epsilon)$  è la soluzione locale del Problema (60) che si ottiene per piccoli valori di  $\epsilon$ , dovrà risultare  $f(x^*(\epsilon)) \leq f(x^*)$ , in quanto la minimizzazione di  $f$  è effettuata su un insieme più grande. Si ha quindi, tenendo conto della (50), scritta con  $\lambda^*$  al posto di  $\mu^*$ :

$$0 \geq \frac{f(x^*(\epsilon)) - f(x^*)}{\epsilon} \simeq -\lambda^*,$$

che implica appunto  $\lambda^* \geq 0$ .

Abbiamo così in pratica dimostrato la proposizione seguente.

**Teorema 28** Sia  $x^*$  un punto di regolarità per i vincoli del Problema (51). Condizione necessaria affinché  $x^*$  sia una soluzione locale del problema è che esista un vettore di moltiplicatori  $\lambda^*$ , con componenti  $\lambda_1^*, \lambda_2^*, \dots, \lambda_p^*$ , tale che:

$$\nabla_x L(x^*, \lambda^*) = \nabla f(x^*) + \sum_{i=1}^p \lambda_i^* \nabla g_i(x^*) = 0, \quad (61)$$

$$\lambda_i^* \geq 0, \quad i = 1, \dots, p \quad (62)$$

$$\lambda_i^* g_i(x^*) = 0, \quad i = 1, \dots, p \quad (63)$$

o, in forma matriciale:

$$\nabla_x L(x^*, \lambda^*) = \nabla f(x^*) + \nabla g(x^*) \lambda^* = 0, \quad (64)$$

$$\lambda^* \geq 0, \quad (65)$$

$$\lambda^{*T} g(x^*) = 0. \quad (66)$$

Le condizioni necessarie di ottimalità della proposizione precedente sono note come *condizioni di Kuhn-Tucker* (o anche, di *Karush-Kuhn-Tucker*). Un punto  $x^*$  viene detto *punto di Kuhn-Tucker*, o sinteticamente punto di KT, se esiste un moltiplicatore  $\lambda^*$  tale che  $x^*, \lambda^*$  soddisfano le condizioni.

La condizione (63) viene chiamata *condizione di complementarità*; notiamo che la (63) implica la (66), mentre il viceversa è vero solo tenendo conto anche della (65) e del fatto che  $g(x^*) \leq 0$ .

Si dice inoltre che un punto di KKT  $x^*$  soddisfa la condizione di **stretta complementarità** se risulta  $\lambda_i^* > 0$  per ogni  $i$ :  $g_i(x^*) = 0$ .

La condizione necessaria  $\nabla_x L(x^*, \mu^*) = 0$ , e quella di complementarità  $\lambda_i^* g_i(x^*) = 0$ ,  $i = 1, \dots, p$  danno luogo ad un sistema di  $n + p$  equazioni nelle  $n + p$  incognite  $x, \lambda$ :

$$\begin{aligned} \nabla_x L(x, \lambda) &= 0, \\ \lambda_i g_i(x) &= 0, \quad i = 1, \dots, p. \end{aligned}$$

Tra le soluzioni  $\bar{x}, \bar{\lambda}$  di questo sistema, per cui risulti anche  $g_i(\bar{x}) \leq 0, \bar{\lambda}_i \geq 0, i = 1, \dots, p$ , si hanno tutte le soluzioni locali regolari del problema (51) e i corrispondenti moltiplicatori.

**Osservazione 3** Osserviamo che, nel caso di vincoli  $g(x)$  siano lineari, la regolarità dei vincoli non è necessaria (come nel caso di vincoli di uguaglianza).

Se  $f(x)$  e  $h(x)$  sono due volte continuamente differenziabile, vale anche la seguente condizione necessaria del secondo ordine, in cui  $\nabla g_a(x^*)$  denota la matrice gradiente dei vincoli attivi in  $x^*$ .

**Teorema 29** Sia  $x^*$  un punto di regolarità per i vincoli del Problema (51). Condizione necessaria affinché  $x^*$  sia una soluzione locale del Problema (51) è che risulti  $d^T \nabla_x^2 L(x^*, \lambda^*) d \geq 0$  per ogni  $d \in \mathbb{R}^n$  tale che  $\nabla g_a(x^*)^T d = 0$ .

Per enunciare anche una condizione sufficiente del secondo ordine, bisogna premettere la definizione di vincolo *strettamente attivo* in un punto di Kuhn-Tucker  $x^*$ . Un vincolo  $g_i(x)$  è strettamente attivo in  $x^*$  se è attivo in  $x^*$ , e se il corrispondente moltiplicatore

$\lambda_i^*$  è strettamente positivo; se denotiamo con  $I_{sa}(x^*)$  l'insieme degli indici dei vincoli di disuguaglianza strettamente attivi in  $x^*$ , abbiamo:

$$I_{sa}(x^*) = \{i \in \{1, \dots, p\} : g_i(x^*) = 0, \lambda_i^* > 0\}.$$

Osserviamo che se il punto  $x^*$  soddisfa la condizione di stretta complementarità, allora  $I_{sa}(x^*) = I_a(x^*)$ .

Possiamo denotare con  $g_{sa}$  il sottovettore di  $g$  costituito dai soli vincoli strettamente attivi in  $x^*$ :  $g_{sa}(x) = [g_i(x)], i \in I_{sa}(x^*)$ , e con  $\nabla g_{sa}$  la matrice gradiente di  $g_{sa}$ . Abbiamo allora la seguente condizione sufficiente del secondo ordine:

**Teorema 30** *Condizione sufficiente affinché un punto ammissibile  $x^*$  sia una soluzione locale stretta del Problema (51) è che esista un moltiplicatore  $\lambda^* \geq 0$  tale che  $\nabla_x L(x^*, \lambda^*) = 0$ ,  $\lambda^{*T} g(x^*) = 0$ , e che risulti  $d^T \nabla_x^2 L(x^*, \lambda^*) d > 0$  per ogni  $d \in \mathbb{R}^n$  tale che  $\nabla g_{sa}(x^*)^T d = 0$ ,  $d \neq 0$ .*

Notiamo che nella condizione sufficiente non è richiesto che  $x^*$  sia un punto di regolarità.

## 7.10 Il caso vincolato: vincoli di disuguaglianza e di uguaglianza

Sulla base dei risultati dei paragrafi precedenti, possiamo facilmente dedurre le condizioni di ottimalità per il caso più generale del Problema (3), in cui sono presenti sia vincoli di disuguaglianza che vincoli di uguaglianza. Riscriviamo per comodità il problema, nella forma scalare:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & g_i(x) \leq 0, \quad i = 1, \dots, p, \\ & h_j(x) = 0, \quad j = 1, \dots, m; \end{aligned} \tag{67}$$

l'associata funzione Lagrangiana è data da:

$$L(x, \lambda) = f(x) + \sum_{i=1}^p \lambda_i g_i(x) + \sum_{j=1}^m \mu_j h_j(x).$$

Abbiamo la seguente condizione necessaria del primo ordine, cui di solito si fa ancora riferimento con il nome di condizioni di Kuhn-Tucker:

**Teorema 31** *Sia  $x^*$  un punto di regolarità per i vincoli del Problema (67). Condizione necessaria affinché  $x^*$  sia una soluzione locale del problema è che esistano un vettore di moltiplicatori  $\lambda^*$ , con componenti  $\lambda_1^*, \dots, \lambda_p^*$ , e un vettore di moltiplicatori  $\mu^*$ , con componenti  $\mu_1^*, \dots, \mu_m^*$ , tali che:*

$$\nabla_x L(x^*, \lambda^*, \mu^*) = \nabla f(x^*) + \sum_{i=1}^p \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^m \mu_j^* \nabla h_j(x^*) = 0, \tag{68}$$

$$\lambda_i^* \geq 0, \quad i = 1, \dots, p \tag{69}$$

$$\lambda_i^* g_i(x^*) = 0, \quad i = 1, \dots, p. \tag{70}$$

La condizione necessaria  $\nabla_x L(x^*, \lambda^*, \mu^*) = 0$ , quella di complementarità  $\lambda_i^* g_i(x^*) = 0$ ,  $i = 1, \dots, p$  e quella di ammissibilità  $h_j(x^*) = 0$ ,  $j = 1, \dots, m$  danno luogo ad un sistema di  $n + p + m$  equazioni nelle  $n + p + m$  incognite  $x, \lambda, \mu$ :

$$\begin{aligned} \nabla_x L(x, \lambda, \mu) &= 0, \\ \lambda_i g_i(x) &= 0, \quad i = 1, \dots, p, \\ h_j(x) &= 0, \quad j = 1, \dots, m. \end{aligned}$$

Tra le soluzioni  $\bar{x}, \bar{\lambda}, \bar{\mu}$  di questo sistema, per cui risulti anche  $g_i(\bar{x}) \leq 0$ ,  $\bar{\lambda}_i \geq 0$ ,  $i = 1, \dots, p$ , si hanno tutte le soluzioni locali regolari del problema (67) e i corrispondenti moltiplicatori. Ovviamente la difficoltà di procedere alla ricerca delle soluzioni di questo sistema è ancora maggiore che nei casi più semplici trattati in precedenza. Non è invece difficile, dato un punto  $x^k$  prodotto alla generica iterazione di un algoritmo, verificare se esistono moltiplicatori  $\lambda^k, \mu^k$  tali che le condizioni siano soddisfatte, essendo le condizioni lineari rispetto ai moltiplicatori. Pertanto nel caso vincolato le condizioni necessarie di Kuhn-Tucker sono alla base della definizione di quell'insieme  $\Omega$  introdotto nel paragrafo 7.1.

Occorre sottolineare che le condizioni di Kuhn-Tucker forniscono solo le soluzioni regolari; un'analisi completa del Problema (67) richiede anche di valutare la funzione obiettivo nei punti ammissibili che non sono di regolarità.

Per quel che riguarda le condizioni del secondo ordine, nell'ipotesi che le funzioni del problema siano due volte continuamente differenziabili, abbiamo le proposizioni seguenti.

**Teorema 32** *Sia  $x^*$  un punto di regolarità per i vincoli del Problema (67). Condizione necessaria affinché  $x^*$  sia una soluzione locale del Problema (67) è che esistano moltiplicatori  $\lambda^* \geq 0$  e  $\mu^*$  tali che  $\nabla_x L(x^*, \lambda^*, \mu^*) = 0$ ,  $\lambda^{*T} g(x^*) = 0$ , e che per ogni  $d \in \mathbb{R}^n$  tale che:*

$$\begin{bmatrix} \nabla g_a(x^*)^T \\ \nabla h(x^*)^T \end{bmatrix} d = 0$$

*risulti*

$$d^T \nabla_x^2 L(x^*, \lambda^*, \mu^*) d \geq 0.$$

**Teorema 33** *Condizione sufficiente affinché un punto ammissibile  $x^*$  sia una soluzione locale stretta del Problema (67) è che esistano moltiplicatori  $\lambda^* \geq 0$  e  $\mu^*$  tali che  $\nabla_x L(x^*, \lambda^*, \mu^*) = 0$ ,  $\lambda^{*T} g(x^*) = 0$ , e che per ogni  $d \in \mathbb{R}^n$  con  $d \neq 0$  e tale che*

$$\begin{bmatrix} \nabla g_{sa}(x^*)^T \\ \nabla h(x^*)^T \end{bmatrix} d = 0$$

*risulti*

$$d^T \nabla_x^2 L(x^*, \lambda^*, \mu^*) d > 0.$$

Notiamo che nella condizione sufficiente compaiono solo i vincoli di disuguaglianza strettamente attivi, e non è richiesto che  $x^*$  sia un punto di regolarità.

**Osservazione 4** *Si osservi che nel Teorema 33, si deve verificare se una condizione sull'hessiano del Lagrangiano per direzioni che si trovano nell'insieme*

$$\mathcal{Y}^* = \{d \in \mathbb{R}^n : d \neq 0 \quad \nabla g_{sa}(x^*)^T d = 0, \quad \nabla h(x^*)^T d = 0\}.$$

*Nel caso in cui  $\mathcal{Y}^* = \emptyset$ , ovvero l'unica direzione che soddisfa il sistema*

$$\nabla g_{sa}(x^*)^T d = 0 \quad \nabla h(x^*)^T d = 0,$$

*è  $d = 0$ , la condizione sufficiente si intende banalmente soddisfatta.*

**Osservazione 5** *Osserviamo infine che, nel caso di vincoli lineari, la regolarità dei vincoli non è richiesta neanche per le condizioni necessarie.*

Come nel caso non vincolato, anche nel caso vincolato le condizioni necessarie del secondo ordine sono di difficile utilizzo. Sono invece di facile verifica le condizioni sufficienti del secondo ordine.

Notiamo infatti che, nei vari problemi vincolati considerati, le condizioni del secondo ordine richiedono, innanzi tutto, che siano soddisfatte le condizioni necessarie del primo ordine, sia pure in assenza di regolarità del punto considerato; e poi che sia definita positiva la matrice delle derivate seconde della funzione Lagrangiana, su un insieme che non è tutto  $\mathbb{R}^n$ , come nel caso non vincolato, bensì un sottospazio lineare dato dalle soluzioni di un sistema di equazioni lineari omogenee nella variabile  $d$ , in cui la matrice dei coefficienti è data dal gradiente trasposto dei vincoli di disuguaglianza strettamente attivi e/o dei vincoli di uguaglianza.

Facendo ad esempio riferimento alla Proposizione 33, se si pone  $M = \nabla_x^2 L(x^*, \lambda^*, \mu^*)$  e  $B = [\nabla g_{sa}(x^*) \nabla h(x^*)]$ , la condizione sufficiente del secondo ordine richiede che  $d^T M d > 0$  per ogni  $d \neq 0 : B^T d = 0$ . Per effettuare questo genere di verifica, possiamo utilizzare il seguente risultato.

**Teorema 34** *Sia  $M$  una matrice quadrata simmetrica  $n \times n$ , sia  $B$  una matrice  $n \times q$ , con  $q < n$ . Risulta  $d^T M d > 0$  per ogni  $d$  tale che  $B^T d = 0$ ,  $d \neq 0$ , se e solo se, considerata la matrice di dimensione  $(q + n) \times (q + n)$ :*

$$\begin{pmatrix} 0 & B^T \\ B & M \end{pmatrix},$$

*gli ultimi  $(n - q)$  minori principali della medesima hanno determinanti il cui segno è pari a quello di  $(-1)^q$ .*

Quindi anche nel caso vincolato la verifica della condizione sufficiente del secondo ordine si riconduce ad un semplice calcolo di determinanti.

**Osservazione 6** *Si osservi che la condizione del Teorema 34, si applica solo quando  $q < n$  ovvero quando il numero dei vincoli di uguaglianza e dei vincoli di disuguaglianza strettamente attivi è minore della dimensione  $n$ .*

Concludiamo con l'osservazione che i moltiplicatori  $\lambda^*$ ,  $\mu^*$ , forniscono i coefficienti di sensibilità del valore ottimo della funzione obiettivo  $f(x^*)$  rispetto alle variazioni dei termini noti dei vincoli, rispettivamente di disuguaglianza e di uguaglianza. Ovviamente, per un vincolo non attivo, il coefficiente di sensibilità risulta nullo.

### 7.11 Il caso convesso

Nella Proposizione 9 abbiamo fatto vedere che un problema di ottimizzazione in cui  $f$  e  $g_i$ ,  $i = 1, \dots, p$  sono funzioni convesse e  $h_j$ ,  $j = 1, \dots, m$  sono funzioni affini, risulta essere un problema di ottimizzazione convesso. Facciamo ora vedere l'importante proprietà che, per un problema di questo tipo, le condizioni di Kuhn-Tucker, anche in assenza di regolarità, risultano essere condizioni sufficienti di ottimo globale.

Riscriviamo per comodità il problema:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ & g_i(x) \leq 0, \quad i = 1, \dots, p, \\ & a_j^T x - b_j = 0, \quad j = 1, \dots, m, \end{aligned} \tag{71}$$

e dimostriamo la proposizione seguente.

**Teorema 35** *Si supponga che, nel Problema (71),  $f(x)$  sia una funzione (strettamente) convessa in  $\mathbb{R}^n$ , e che  $g_i(x)$  siano funzioni convesse in  $\mathbb{R}^n$ . Sia  $x^*$  un punto ammissibile per il problema, e si supponga che esistano moltiplicatori  $\lambda_i^*$  e  $\mu_j^*$  tali che:*

$$\nabla_x L(x^*, \lambda^*, \mu^*) = \nabla f(x^*) + \sum_{i=1}^p \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^m \mu_j^* a_j = 0, \quad (72)$$

$$\lambda_i^* \geq 0, \quad i = 1, \dots, p \quad (73)$$

$$\lambda_i^* g_i(x^*) = 0, \quad i = 1, \dots, p. \quad (74)$$

Allora  $x^*$  è una soluzione globale (stretta) per il Problema (71).

**Dimostrazione.** Sia  $\bar{x}$  un qualunque punto ammissibile per il problema. Risulta  $g_i(\bar{x}) \leq 0$ , e quindi  $\lambda_i^* g_i(\bar{x}) \leq 0$ ; inoltre risulta  $a_j^T \bar{x} - b_j = 0$ . Possiamo allora scrivere la disuguaglianza:

$$f(\bar{x}) \geq f(x^*) + \sum_{i=1}^p \lambda_i^* g_i(\bar{x}) + \sum_{j=1}^m \mu_j^* (a_j^T \bar{x} - b_j). \quad (75)$$

Per la convessità di  $f$  e  $g_i$  abbiamo che

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

$$g_i(\bar{x}) \geq g_i(x^*) + \nabla g_i(x^*)^T (\bar{x} - x^*).$$

Sostituendo nella (75) e notando che  $a_j^T \bar{x} - b_j = a_j^T (\bar{x} - x^*)$ , otteniamo:

$$f(\bar{x}) \geq f(x^*) + \nabla f(x^*)^T (\bar{x} - x^*) + \sum_{i=1}^p \lambda_i^* (g_i(x^*) + \nabla g_i(x^*)^T (\bar{x} - x^*)) + \sum_{j=1}^m \mu_j^* a_j^T (\bar{x} - x^*). \quad (76)$$

Poichè  $a_j$  risulta essere il gradiente del vincolo lineare, la precedente disuguaglianza può essere riscritta, mettendo in evidenza  $(\bar{x} - x^*)$ :

$$f(\bar{x}) \geq f(x^*) + \nabla_x L(x^*, \lambda^*, \mu^*)^T (\bar{x} - x^*) + \sum_{i=1}^p \lambda_i^* g_i(x^*); \quad (77)$$

da cui si ottiene, tenendo conto delle condizioni di KT  $\nabla_x L(x^*, \lambda^*, \mu^*) = 0$ ,  $\lambda_i^* g_i(x^*) = 0$ ,  $i = 1, \dots, p$ , che  $f(\bar{x}) \geq f(x^*)$ . Poichè  $\bar{x}$  è un qualunque punto ammissibile per il Problema (71), risulta che  $x^*$  è una soluzione globale. Si lascia per esercizio la conclusione che, nel caso in cui  $f$  è strettamente convessa,  $x^*$  è una soluzione globale stretta.  $\square$

Osserviamo quindi che, nel caso convesso, le condizioni di KKT sono necessarie sotto ipotesi di regolarità dei vincoli, e sufficienti senza ipotesi di regolarità.

Nel caso particolare di problemi con vincoli lineari, in cui quindi l'insieme ammissibile è un poliedro, la regolarità dei vincoli non è richiesta nemmeno per la parte necessarie. Possiamo quindi enunciare il seguente risultato:

**Teorema 36** *Si supponga che, nel Problema (67),  $f(x)$  sia una funzione (strettamente) convessa in  $\mathbb{R}^n$ , e i vincoli di uguaglianza e disuguaglianza siano lineari. Condizione necessaria e sufficiente affinché un punto  $x^*$  ammissibile per il problema sia un minimo globale (stretto) è che valgano le condizioni di KKT.*

## 8 La dualità nella Programmazione Lineare

Il problema di Programmazione Lineare è un problema di ottimizzazione convesso. Inoltre abbiamo affermato che nel caso di vincoli lineari le condizioni necessarie di ottimo valgono anche in assenza dell'ipotesi di regolarità. Quindi vale il risultato enunciato nel Teorema 36.

Le condizioni necessarie e sufficienti di ottimalità consentono di dedurre alcuni risultati molto importanti nella Teoria della Programmazione Lineare che sono noti sotto il nome di Teoria della Dualità.

Facciamo riferimento per semplicità ad un problema di Programmazione Lineare del tipo

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ & Ax \geq b \\ & x \geq 0, \end{aligned} \tag{78}$$

ove  $A$  è una matrice  $p \times n$ ,  $c$  è un vettore di dimensione  $n$ ,  $b$  è un vettore di dimensione  $p$ . Sappiamo che le condizioni di KKT sono necessarie e sufficienti di ottimo globale. Quindi una soluzione del problema (78) esiste se e solo se esistono dei moltiplicatori che soddisfano le condizioni di KKT.

Consideriamo ora il problema di Programmazione Lineare

$$\begin{aligned} \max_{u \in \mathbb{R}^p} \quad & b^T u \\ & A^T u \leq c \\ & u \geq 0; \end{aligned} \tag{79}$$

Anche in questo caso le condizioni di KKT sono necessarie e sufficienti di ottimo globale. La coppia di problemi di programmazione lineare (78) e (79) è detta coppia *primale-duale simmetrica*.

Si può dimostrare che le condizioni di KKT dei due problemi di programmazione lineare (78) e (79) sono le stesse. Si tratta del teorema della dualità forte. Il teorema della dualità forte fornisce quindi una relazione tra le soluzioni ottime di questi due problemi.

**Teorema 37 (Teorema della Dualità Forte)** *Il problema primale (78) ha una soluzione ottima  $x^*$  se e solo se il problema duale (79) ha una soluzione ottima  $u^*$ ; inoltre i valori delle funzioni obiettivo coincidono all'ottimo  $c^T x^* = b^T u^*$ .*

**Dimostrazione.** Si tratta di fare vedere che le condizioni necessarie e sufficienti di ottimo per il problema (78) risultano essere necessarie e sufficienti di ottimo globale  $u^*$  per il problema (79).

Chiamiamo con  $u \in \mathbb{R}^p$  il moltiplicatore associato ai vincoli  $Ax \geq b$ , e con  $v \in \mathbb{R}^n$  il moltiplicatore associato ai vincoli  $x \geq 0$ ; con queste notazioni, la funzione Lagrangiana del problema (78) è data da:

$$L_p(x, u, v) = c^T x + u^T(b - Ax) - v^T x,$$

e le condizioni di KT affermano l'esistenza di  $u^* \geq 0, v^* \geq 0$  tali che:

$$\nabla_x L_p(x^*, u^*, v^*) = c - A^T u^* - v^* = 0, \tag{80}$$

$$u^{*T}(b - Ax^*) = 0 \tag{81}$$

$$v^{*T} x^* = 0 \tag{82}$$

È possibile riscrivere queste condizioni eliminando  $v^*$ . Infatti dalla (80) otteniamo  $c - A^T u^* = v^*$  ed imponendo la condizione  $v^* \geq 0$ , si ottiene

$$c - A^T u^* \geq 0.$$

Sostituendo l'espressione della  $v^*$  nella (82) otteniamo poi  $(c - A^T u^*)^T x^* = 0$ , ovvero

$$c^T x^* = u^{*T} A x^*.$$

Sviluppando la (81) si ha anche  $u^{*T} b - u^{*T} A x^* = 0$ , che insieme alla precedente consente di scrivere

$$u^{*T} b = u^{*T} A x^* = c^T x^*.$$

Possiamo quindi affermare che,  $x^*$  è una soluzione del problema (78) se e solo se esiste un moltiplicatore  $u^* \in \mathbb{R}^m$  tale che  $x^*$  e  $u^*$  soddisfano le condizioni:

$$A x^* \geq b, \quad x^* \geq 0, \quad (83)$$

$$A^T u^* \leq c, \quad u^* \geq 0. \quad (84)$$

$$c^T x^* = b^T u^* \quad (85)$$

Osserviamo che alla condizione di ammissibilità di  $x^*$  (83), si aggiungono una condizione nella sola  $u^*$  data dalla (84) e una condizione che coinvolge sia  $x^*$  che  $u^*$ . In particolare la condizione (84) evidenzia che  $u^* \in \mathbb{R}^m$  è ammissibile per il problema duale (79).

Scriviamo ora le condizioni di KKT per il problema (79). Chiamiamo con  $x \in \mathbb{R}^n$  il moltiplicatore associato ai vincoli  $A^T u \leq c$ , e con  $z \in \mathbb{R}^p$  il moltiplicatore associato ai vincoli  $u \geq 0$ ; con queste notazioni, osservando che il problema duale è un problema di massimizzazione, abbiamo per il problema la funzione Lagrangiana:

$$L_d(u, x, z) = -b^T u + x^T (A^T u - c) - z^T u,$$

e se  $u^*$  è una soluzione del problema duale, le condizioni di KT affermano l'esistenza di  $x^* \geq 0, z^* \geq 0$  tali che:

$$\nabla_u L_d(u^*, x^*, z^*) = -b + A x^* - z^* = 0, \quad (86)$$

$$x^{*T} (A^T u^* - c) = 0 \quad (87)$$

$$z^{*T} u^* = 0 \quad (88)$$

In modo analogo a quanto fatto sul problema primale, possiamo eliminare  $z^*$ . Infatti dalla (86) otteniamo  $-b + A x^* = z^* \geq 0$ ; premoltiplicando la (86) per  $u^{*T}$ , e tenendo conto delle (87), (88), otteniamo poi  $-b^T u^* + c^T x^* = 0$ . Possiamo quindi affermare che, se  $u^*$  è una soluzione del problema duale, esiste un moltiplicatore  $x^*$  tale che  $u^*$  e  $x^*$  soddisfano le condizioni:

$$A^T u^* \leq c, \quad u^* \geq 0,$$

$$A x^* \geq b \quad x^* \geq 0,$$

$$b^T u^* = c^T x^*.$$

Si constata che le condizioni per il primale (78) e quelle per il duale (79) coincidono. Ricordando che nel caso convesso le condizioni necessarie sono anche sufficienti, possiamo

concludere quanto segue: se il primale ha una soluzione  $x^*$  esiste un moltiplicatore  $u^*$  che, insieme ad  $x^*$ , soddisfa le condizioni necessarie per il primale; ma poichè gli stessi  $u^*$  e  $x^*$  soddisfano le condizioni sufficienti per il duale, si deduce che il duale ha la soluzione  $u^*$ , con moltiplicatore  $x^*$ ; viceversa, se il duale ha una soluzione  $u^*$ , con associato moltiplicatore  $x^*$ , allora  $x^*$  è soluzione del primale, con associato moltiplicatore  $u^*$ . Inoltre il valore ottimo della funzione obiettivo primale coincide con il valore ottimo della funzione obiettivo duale, come risulta dalla (85).  $\square$

In conclusione, da un punto di vista analitico, risolvere il primale è del tutto equivalente a risolvere il duale. Non è però così da un punto di vista algoritmico, in quanto i due problemi sono diversi, e, a seconda dell'algoritmo che si utilizza, può essere più conveniente risolvere un problema piuttosto che l'altro.

Si può dimostrare anche un altro importante risultato della teoria della dualità che è noto con il nome di *Teorema della Dualità Debole*:

**Teorema 38 (Teorema della Dualità Debole)** *Siano  $\bar{x}$  e  $\bar{u}$  soluzioni ammissibile rispettivamente per il problema primale e per il problema duale. Allora risulta  $c^T \bar{x} \geq b^T \bar{u}$ .*

**Dimostrazione.** Supponiamo di avere due punti  $\bar{x}$  e  $\bar{u}$  ammissibili rispettivamente per il problema primale e per il problema duale, ovvero tali che:

$$\begin{aligned} A^T \bar{u} &\leq c, & \bar{u} &\geq 0, \\ A \bar{x} &\geq b & \bar{x} &\geq 0. \end{aligned}$$

Osserviamo che  $\bar{u}^T (b - A \bar{x}) = \sum_{i=1}^p \bar{u}_j (b_i - a_i^T \bar{x}) \leq 0$  e quindi possiamo scrivere

$$c^T \bar{x} \geq c^T \bar{x} + \bar{u}^T (b - A \bar{x}) = b^T \bar{u} + (c^T - \bar{u}^T A) \bar{x} \geq b^T \bar{u},$$

dove l'ultima disuguaglianza segue dal fatto che  $(c^T - \bar{u}^T A) \bar{x} \geq 0$ .  $\square$

Osserviamo che, se  $x^*$  è una soluzione ottima per il primale (e quindi in particolare ammissibile), e se  $\bar{u}$  è un punto ammissibile per il duale, abbiamo dal teorema della Dualità Debole che

$$c^T x^* \geq b^T \bar{u}.$$

Osserviamo inoltre che, se  $\bar{x}$  e  $x^*$  sono rispettivamente un punto ammissibile e una soluzione per il primale, si ha ovviamente

$$c^T \bar{x} \geq c^T x^*.$$

Quindi un punto ammissibile per il primale, e un punto ammissibile per il duale forniscono, rispettivamente, una limitazione superiore e una limitazione inferiore per il valore ottimo del primale, cioè

$$b^T \bar{u} \leq c^T x^* \leq c^T \bar{x}.$$

Sapere che il valore ottimo  $c^T x^*$  del primale di un problema di Programmazione Lineare si colloca nell'intervallo  $[b^T \bar{u}, c^T \bar{x}]$ , ove  $\bar{x}$  è un qualunque punto ammissibile per il primale, e  $\bar{u}$  è un qualunque punto ammissibile per il duale, risulta di grande utilità dal punto di vista algoritmico.

Naturalmente si può dare un risultato analogo per la soluzione ottima del duale  $u^*$ . In particolare se  $\bar{x}$  e  $\bar{u}$  soluzioni ammissibile rispettivamente per il problema primale e per il problema duale, possiamo scrivere

$$b^T \bar{u} \leq b^T u^* \leq c^T \bar{x}.$$

Fino ad ora abbiamo fatto riferimento ad una coppia primale -duale simmetrica. Si possono dare delle regole generali per scrivere il problema duale di un qualunque problema di programmazione lineare. Tali regole sono schematizzate nella tabella, in cui, data una matrice  $A$ , indichiamo con  $a_i^T$  le righe di  $A$  e con  $A_j$  le colonne di  $A$ .

PRIMALE	DUALE
$\min c^T x$ $a_i^T x \geq b_i$ $a_i^T x = b_i$ $x_j \geq 0$ $x_j$ non vincolata in segno	$\max b^T u$ $u_i \geq 0$ $u_i$ non vincolata in segno $u^T A_j \leq c_j$ $u^T A_j = c_j$
DUALE	PRIMALE

**Osservazione 7** Osserviamo che il ruolo di primale e di duale sono intercambiabili. Inoltre facendo il duale del problema duale si riottiene il problema primale.

**Osservazione 8** Nel derivare i risultati della Teoria della dualità abbiamo fatto riferimento alla coppia prima -duale simmetrica. Questi risultati sono validi comunque si scelga una coppia primale-duale, facendo attenzione al segno delle disuguaglianze, in particolare nel teorema della dualità debole, che sono date facendo riferimento ad un problema primale di minimizzazione ed uno duale di massimizzazione.

Nella successiva tabella riportiamo le situazioni che si possono verificare (Sì) e quelle che non si possono verificare (No) in una coppia primale-duale:

		DUALE		
		ottimo finito	illimitato superior.	ins. ammissibile vuoto
PRIMALE	ottimo finito	Sì	No	No
	illimitato inferior.	No	No	Sì
	ins. ammissibile vuoto	No	Sì	Sì

**Esempio 13** Sia dato il problema di Programmazione Lineare

$$\begin{aligned} \min \quad & x_1 - 2x_2 \\ & -x_1 + x_2 \leq -3 \\ & x_1 + x_2 \leq 5 \\ & x_1 \geq 0, \quad x_2 \geq 0 \end{aligned}$$

1. Scrivere il problema duale;
2. dare una limitazione superiore ed inferiore al valore ottimo  $c^T x^*$  del problema primale;
3. dare una limitazione superiore ed inferiore al valore ottimo  $b^T u^*$  del problema duale;
4. utilizzando il teorema della dualità forte (o le equivalenti condizioni di KKT) dire se il punto  $x^* = (4, 1)^T$  è ottimo per il primale e fornire a certificato il valore ottimo della soluzione duale  $u^*$ .

Il problema duale è (attenzione per applicare le regole della tabella è necessario che i problemi siano scritti in una delle forme previste):

$$\begin{aligned} \max \quad & 3u_1 - 5u_2 \\ & u_1 - u_2 \leq 1 \\ & -u_1 - u_2 \leq -2 \\ & u_1 \geq 0, \quad u_2 \geq 0 \end{aligned}$$

Si verifica facilmente che il punto  $\bar{x} = (3, 0)^T$  è ammissibile per il primale, mentre il punto  $\bar{u} = (0, 2)^T$  è ammissibile per il duale. Si ottengono quindi i due intervalli

$$\begin{aligned} -10 = b^T \bar{u} &\leq c^T x^* \leq c^T \bar{x} = 3 \\ -10 = b^T \bar{u} &\leq b^T u^* \leq c^T \bar{x} = 3. \end{aligned}$$

Le condizioni di KKT (equivalenti al teorema della dualità forte, si scrivono;

$$\begin{aligned} -x_1 + x_2 &\leq -3 \\ x_1 + x_2 &\leq 5 \\ x_1 &\geq 0, \quad x_2 \geq 0 \\ u_1 - u_2 &\leq 1 \\ -u_1 - u_2 &\leq -2 \\ u_1 &\geq 0, \quad u_2 \geq 0 \\ x_1(u_1 - u_2 - 1) &= 0 \\ x_2(-u_1 - u_2 + 2) &= 0 \\ u_1(-x_1 + x_2 + 3) &= 0 \\ u_2(x_1 + x_2 - 5) &= 0 \end{aligned}$$

Il punto  $x^*$  è ammissibile per il primale. Dalle condizioni di complementarità segue anche che per essere ottimo deve esistere un  $u^* \geq 0$  tale che  $u_1^* - u_2^* - 1 = 0$  e  $-u_1^* - u_2^* + 2 = 0$ , da cui si ottiene  $u^* = (3/2, 1/2)$ .

## 9 Generalità sugli algoritmi di ottimizzazione

I problemi di ottimizzazione che si presentano nella pratica sono di solito così complessi che non è possibile determinarne una soluzione per via analitica. La complessità è determinata innanzi tutto dal numero di variabili e di vincoli, che definiscono la *dimensione* del problema; e poi dalla eventuale presenza di funzioni non lineari tra le funzioni  $f, g_i, h_j$ . La soluzione analitica è possibile solo nel caso di poche variabili e di funzioni estremamente semplici, e cioè solo nei casi che si utilizzano come esempi ed esercizi nei testi e sulla lavagna.

Nella pratica, per risolvere un problema di ottimizzazione occorre fare ricorso ad un *algoritmo iterativo*, cioè ad un programma di calcolo che, data una approssimazione corrente  $x^k$  della soluzione, determina, con una appropriata sequenza di operazioni, una nuova approssimazione  $x^{k+1}$ . A partire da una approssimazione iniziale  $x^0$  si determina così una successione  $\{x^k\}$ .

Occorre però a questo punto mettere in evidenza una limitazione intrinseca degli algoritmi di ottimizzazione, che consiste nel fatto che, per come sono costruiti, sono in grado di determinare solo punti che, per un dato problema, ne soddisfano le condizioni necessarie di ottimalità: cioè solo punti dell'insieme  $\Omega$  introdotto nel §7.1. Se si denota con  $\mathcal{X}$  l'insieme delle soluzioni locali del problema, risulta evidentemente  $\mathcal{X} \subseteq \Omega \subseteq \mathcal{F}$ . Le prestazioni di un algoritmo vanno perciò valutate in relazione alla sua capacità di determinare punti di  $\Omega$ , che, in questo contesto, viene detto *insieme bersaglio*, piuttosto che in relazione alla capacità di determinare punti di  $\mathcal{X}$ .

Con *convergenza* dell'algoritmo si intende appunto la sua capacità di centrare, con la successione  $\{x^k\}$  che genera, l'insieme bersaglio  $\Omega$ . L'algoritmo si dice *convergente* se fornisce un punto di  $\Omega$  dopo un numero finito di iterazioni, o almeno, al limite, per  $k \rightarrow \infty$ . Nel primo caso si parla di convergenza *finita*, nel secondo di convergenza *asintotica*.

La convergenza finita si consegue solo per problemi particolari e algoritmi specifici per questi; ad esempio, l'algoritmo del simplesso per la Programmazione Lineare ha convergenza finita. Gli algoritmi per la Programmazione Nonlineare hanno in generale convergenza asintotica.

Ovviamente, per un algoritmo con convergenza asintotica, non sarà possibile in pratica eseguire un numero infinito di iterazioni, e occorre quindi prevedere una *criterio d'arresto* e cioè una regola che interrompa l'esecuzione dell'algoritmo dopo un numero finito di iterazioni. Il criterio d'arresto di solito si basa sul riconoscere di avere trovato, se non proprio un punto di  $\Omega$ , almeno una sua buona approssimazione. Ad esempio, per il Problema (2), sappiamo che  $\Omega = \{\omega \in \mathbb{R}^n : \nabla f(\omega) = 0\}$ ; quindi dato un  $\epsilon$  sufficientemente piccolo, si può pensare di arrestare l'algoritmo al primo valore di  $k$  per cui risulti  $\|\nabla f(x^k)\| < \epsilon$ .

Nell'ambito della convergenza asintotica, si possono avere differenti caratterizzazioni della convergenza dell'algoritmo, a seconda del comportamento della successione  $\{x^k\}$  per  $k \rightarrow \infty$ . Un primo caso si ha se l'intera successione è *convergente* a un punto  $\omega \in \Omega$ ; cioè se  $\lim_{k \rightarrow \infty} x^k = \omega \in \Omega$ . Un tale comportamento, per quanto auspicabile, non sempre è realizzabile. Un secondo caso si ha quando la successione  $\{x^k\}$  ha almeno un *punto limite* (o punto di accumulazione)  $\omega \in \Omega$ , e cioè quando è possibile estrarre dalla successione almeno una sottosuccessione (anche detta successione estratta)  $x^{k_j}, j = 1, 2, \dots$  che risulti convergente a un punto  $\omega \in \Omega$ .

Poiché se una successione converge ad un punto  $\omega \in \Omega$  anche tutte le sue sottosuccessioni devono convergere allo stesso punto  $\omega$ , è evidente che nel primo caso l'algoritmo deve soddisfare requisiti molto più stringenti che nel secondo. Tuttavia, nella pratica, anche

il secondo comportamento è più che soddisfacente; infatti, tenuto conto di quanto detto a proposito del criterio d'arresto, si ha che, se  $\{x^k\}$  ha almeno un punto limite in  $\Omega$ , è sempre possibile trovare, per  $k$  sufficientemente grande, un punto  $x^k$  che approssimi un punto di  $\Omega$  con la precisione desiderata.

Ovviamente la successione  $\{x^k\}$  può poi avere più di un punto limite, sia in  $\Omega$  che fuori di  $\Omega$ .

Un'altra caratterizzazione della convergenza di un algoritmo si ha in relazione alla scelta di  $x^0$ . Per alcuni algoritmi, la convergenza ad un punto  $\omega \in \Omega$  si consegue qualunque sia  $x^0 \in \mathbb{R}^n$ ; in questo caso si dice che l'algoritmo ha convergenza *globale*. Per altri algoritmi, la convergenza ad un punto  $\omega \in \Omega$  è assicurata solo se  $x^0 \in \mathcal{S}$ , essendo  $\mathcal{S}$  un intorno sferico aperto di  $\omega$ ; in questo caso si dice che l'algoritmo ha convergenza *locale*. Notiamo che quando si parla di convergenza globale (locale) di un algoritmo, *non si intende* convergenza ad una soluzione globale (locale) del problema.

Si consideri ad esempio il semplice problema:

$$\min_{x \in \mathbb{R}} x^2 \quad (89)$$

la cui soluzione è ovviamente  $x^* = 0$ .

Un algoritmo per risolverlo può essere dato dal procedimento iterativo Alg1 definito da

$$\text{Alg1 : } x^{k+1} = \frac{x^k}{2}$$

A partire da un  $x^0 \in \mathbb{R}$ , otteniamo:

$$x^1 = x^0/2,$$

$$x^2 = x^1/2 = x^0/4,$$

$$x^3 = x^2/2 = x^0/8,$$

$$x^4 = x^3/2 = x^0/16,$$

e così via; quindi, per un  $k$  generico, abbiamo  $x^k = x^0/2^k$ . È evidente che la successione che si ottiene converge a  $x^* = 0$ , soluzione del problema, qualunque sia  $x^0$ ; quindi Alg1 converge, e converge globalmente.

Un secondo algoritmo per risolvere il problema dell'esempio può essere dato dal procedimento iterativo Alg2 definito da

$$\text{Alg2 : } x^{k+1} = \frac{x^k}{(k+1)}$$

A partire da un  $x^0 \in \mathbb{R}$ , otteniamo:

$$x^1 = x^0,$$

$$x^2 = x^1/2 = x^0/2,$$

$$x^3 = x^2/3 = x^0/(2 \times 3),$$

$$x^4 = x^3/4 = x^0/(2 \times 3 \times 4),$$

e così via; quindi, per un  $k$  generico, abbiamo  $x^k = x^0/k!$ . Anche in questo caso, è evidente che la successione che si ottiene converge globalmente a  $x^* = 0$ , soluzione del problema.

Consideriamo ora, sempre per il problema dell'esempio, un terzo algoritmo Alg3, definito dall'iterazione

$$\text{Alg3 : } x^{k+1} = (x^k)^2$$

A partire da  $x^0 \in \mathbb{R}$ , otteniamo:

$$\begin{aligned}x^1 &= (x^0)^2, \\x^2 &= (x^1)^2 = (x^0)^4, \\x^3 &= (x^2)^2 = (x^0)^8, \\x^4 &= (x^3)^2 = (x^0)^{16},\end{aligned}$$

e così via; quindi, per un  $k$  generico, abbiamo  $x^k = (x^0)^{2^k}$ . In questo caso, se assumiamo  $x^0 = 1$ , otteniamo  $x^k = 1$  per ogni  $k$ . Quindi la successione che si ottiene è convergente, ma ad un punto  $x^* = 1 \notin \Omega$ . Se assumiamo  $x^0 = 2$  la successione che si ottiene è evidentemente divergente. Se si assume  $x^0 = 1/2$ , si ottiene  $x^k = \frac{1}{2^{2^k}}$ , e in questo caso la successione che si ottiene converge a  $x^* = 0$ . Si può facilmente verificare che la successione converge alla soluzione del problema  $x^* = 0$  se, e solo se,  $x^0 \in (-1, 1)$ . Quindi Alg3 è un algoritmo che converge solo localmente.

Nel valutare le prestazioni di un algoritmo, oltre alla convergenza, occorre poi prendere in considerazione la sua *rapidità di convergenza* (velocità di convergenza). Infatti dati due algoritmi che risolvono lo stesso problema, il numero di iterazioni richiesto affinché il criterio di arresto sia soddisfatto può risultare molto diverso.

Nella valutazione della rapidità di convergenza si assume, eventualmente estraendo una sottosuccessione convergente ad un punto limite  $\omega \in \Omega$ , che la successione  $\{x^k\}$  prodotta dall'algoritmo risulti convergente ad  $\omega \in \Omega$ ; cosicché risulti  $\lim_{k \rightarrow \infty} x^k = \omega$ , e quindi  $\lim_{k \rightarrow \infty} \|x^k - \omega\| = 0$ .

Il metodo più utilizzato per valutare la rapidità di convergenza ad  $\omega$  della successione  $\{x^k\}$  si basa sull'analisi di come si riduce il termine  $\|x^{k+1} - \omega\|$  rispetto al termine  $\|x^k - \omega\|$ ; si basa cioè sull'analisi del rapporto  $\|x^{k+1} - \omega\|/\|x^k - \omega\|$ .

Si dice che l'algoritmo ha rapidità di convergenza:

- *lineare*, se esiste un  $c \in (0, 1)$  tale che, per  $k$  abbastanza grande, si ha:

$$\frac{\|x^{k+1} - \omega\|}{\|x^k - \omega\|} \leq c;$$

- *superlineare*, se si ha:

$$\lim_{k \rightarrow \infty} \frac{\|x^{k+1} - \omega\|}{\|x^k - \omega\|} = 0;$$

- *quadratica*, se esiste un  $c > 0$  tale che, per  $k$  abbastanza grande, si ha:

$$\frac{\|x^{k+1} - \omega\|}{\|x^k - \omega\|^2} \leq c.$$

Si verifica facilmente che se un algoritmo converge superlinearmente, converge anche linearmente; e se converge quadraticamente converge anche superlinearmente. Quindi la convergenza quadratica è più rapida di quella superlineare, e la superlineare è più rapida della lineare.

Se consideriamo di nuovo il Problema (89), ricordando che  $\omega = 0$  vediamo che:

- per Alg1 si ha:

$$\frac{\|x^{k+1} - \omega\|}{\|x^k - \omega\|} = \frac{x^0/2^{k+1}}{x^0/2^k} = \frac{1}{2},$$

e quindi Alg1 converge linearmente;

- per Alg2 si ha:

$$\lim_{k \rightarrow \infty} \frac{\|x^{k+1} - \omega\|}{\|x^k - \omega\|} = \lim_{k \rightarrow \infty} \frac{x^0 / (k+1)!}{x^0 / k!} = \lim_{k \rightarrow \infty} \frac{1}{k+1} = 0,$$

e quindi Alg2 converge superlinearmente;

- per Alg3 si ha, direttamente dalla definizione:

$$\frac{\|x^{k+1} - \omega\|}{\|x^k - \omega\|^2} = \frac{x^{k+1}}{(x^k)^2} = 1,$$

e quindi Alg3 converge quadraticamente.

Per  $k = 1, 2, \dots, 5$  a partire dallo stesso  $x^0 = 1/2$ , otteniamo con i tre algoritmi i valori di  $x^k$  riportati nella Tabella 1. Se poi poniamo  $k = 10$  risulta, rispettivamente per Alg1, Alg2, Alg3:  $x^{10} = 4.88 \times 10^{-4}$ ,  $x^{10} = 1.37 \times 10^{-7}$ ,  $x^{10} = 5,56 \times 10^{-309}$  (!). Quindi si constata che Alg3 è più veloce di Alg2, e Alg2 è più veloce di Alg1.

$x^0 = \frac{1}{2}$	$x^1$	$x^2$	$x^3$	$x^4$	$x^5$
Alg1	0.25	0.125	0.0625	0.03215	0.015625
Alg2	0.5	0.25	0.08333	0.020833	0.004166
Alg3	0.25	0.0625	0.0039	$1.5 \times 10^{-5}$	$2.32 \times 10^{-9}$

Table 1: rapidità di convergenza

Osserviamo a questo punto che un algoritmo di ottimizzazione deve essere in grado di risolvere non un problema particolare, come il Problema (89), ma una intera classe di problemi: ad esempio un algoritmo di ottimizzazione non vincolata deve potersi applicare a tutti i problemi della forma (2), qualunque sia la funzione  $f(x)$  continuamente differenziabile; e le proprietà di convergenza e rapidità di convergenza dell'algoritmo debbono valere per l'intera classe. Gli algoritmi Alg1, Alg2, Alg3 sono stati introdotti a titolo puramente esemplificativo con riferimento al Problema (89), ma nessuno di essi può essere considerato un algoritmo di ottimizzazione: infatti nessuno di essi è in grado di risolvere il problema  $\min_{x \in \mathbb{R}} (x-1)^2$ , che è evidentemente dello stesso tipo del Problema (89).

Osserviamo inoltre che, da un punto di vista pratico, le considerazioni analitiche sulla rapidità di convergenza non sono sufficienti a stabilire che un algoritmo ha prestazioni migliori di un altro, in quanto tengono conto solo del numero delle iterazioni, e non di quanto ogni iterazione sia onerosa da un punto di vista computazionale. Come vedremo nel seguito, in generale la singola iterazione di un algoritmo con rapidità lineare richiede calcoli molto più semplici di quelli richiesti da un algoritmo con rapidità quadratica. In pratica quindi occorre cercare un compromesso conveniente, tra l'esigenza di non effettuare troppe iterazioni, e quella di avere iterazioni non troppo onerose.

### Domande

1) Sia dato il problema:  $\min_{x \in \mathbb{R}} \sin \pi x$ ; e l'algoritmo Alg definito da:

$$\text{Alg} : x^{k+1} = \begin{cases} \frac{x^k}{2} & k = 0, 2, 4 \dots (\text{interi pari}), \\ 1 - \frac{x^k}{2} & k = 1, 3, 5 \dots (\text{interi dispari}) \end{cases}$$

Sai analizzarne le caratteristiche di convergenza? E di rapidità di convergenza?

2) Sai verificare che un algoritmo che converge superlinearmente converge anche linearmente? E che un algoritmo che converge quadraticamente converge anche superlinearmente?

3) Per il problema dell'esempio (89), ammettiamo come criterio di arresto che per la derivata della funzione obiettivo risulti  $|df(x^k)/dx| \leq 10^{-6}$ . Assumendo  $x^0 = 1/2$ , dopo quante iterazioni si arresteranno Alg1, Alg2 e Alg3?

4) Nella pratica implementazione degli algoritmi, occorre sempre dare un numero massimo di iterazioni  $K$ , raggiunto il quale l'algoritmo viene comunque arrestato. Per il problema dell'esempio (89), ammettiamo dunque che l'arresto avvenga o se  $|df(x^k)/dx| \leq 10^{-12}$ , o se  $k = K = 30$ . Assumendo  $x^0 = 1/2$ , quale delle due condizioni si verificherà prima?

## 10 Algoritmi per l'ottimizzazione non vincolata

### 10.1 Introduzione

Consideriamo il problema di ottimizzazione non vincolata:

$$\min_{x \in R^n} f(x) \tag{90}$$

in cui  $f : R^n \rightarrow R$ . Nel seguito supporremo sempre valida la seguente ipotesi:

**Ipotesi 1** *La funzione  $f : R^n \rightarrow R$  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.*

È noto che, sotto l'ipotesi 1, il Problema (90) ammette un punto di minimo globale  $x^* \in \mathcal{L}_{x^0}$  e che ogni punto di minimo locale di  $f$  in  $\mathcal{L}_{x^0}$  è un *punto stazionario* di  $f$ , ossia un punto in cui si annulla il gradiente.

Gli algoritmi che prenderemo in considerazione consentono, in generale, soltanto la determinazione di punti stazionari di  $f$ , ossia di punti dell'insieme bersaglio  $\Omega := \{\omega \in R^n : \nabla f(\omega) = 0\}$ . In generale si riesce a garantire che, se  $x^0$  non è già un punto stazionario, vengano comunque ottenuti punti stazionari 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 poi 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.

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 ricerca*  $d^k \in R^n$ .
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.

Dal punto di vista geometrico l'algoritmo si può descrivere come una successione di spostamenti (definiti dagli scalari  $\alpha^k$ ) lungo le direzioni di ricerca  $d^k$ , effettuati a partire da  $x^0$ .

Gli algoritmi che consideriamo sono algoritmi *di discesa* ovvero garantiscono che

$$f(x^k + \alpha^k d^k) < f(x^k). \quad (91)$$

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  $x^0$  e quindi siamo interessati a definire algoritmi le cui proprietà di convergenza siano indipendenti dalla scelta del punto iniziale (algoritmo *globalmente* convergente).

Nella soluzione di problemi applicativi 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  $\Omega$  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 prima possibilità consiste nell'arrestare l'algoritmo quando

$$\|\nabla f(x^k)\| \leq \varepsilon_1 \quad (92)$$

in cui  $\varepsilon_1 > 0$  è un valore sufficientemente piccolo.

Dal punto di vista numerico tale criterio può non essere del tutto soddisfacente perché non fa riferimento né alla precisione del mezzo di calcolo, né alla scala con cui è calcolato  $\nabla f$ . Nei codici di calcolo occorrerà quindi definire criteri più significativi. Osserviamo che la possibilità di utilizzare la (92) (opportunosamente rielaborata) come criterio di arresto richiede che si possa dimostrare che l'algoritmo consente

di raggiungere valori arbitrariamente piccoli di  $\|\nabla f(x^k)\|$  per valori sufficientemente elevati di  $k$ .

Accanto alla condizione (92) se ne utilizzano altre

$$\|x^{k+1} - x^k\| \leq \varepsilon_2 \quad |f(x^{k+1}) - f(x^k)| \leq \varepsilon_3$$

L'uso della condizione  $\|x^{k+1} - x^k\| \leq \varepsilon_2$  è motivata dal fatto che nel caso di convergenza superlineare la quantità  $\|x^{k+1} - x^k\|$  fornisce una stima dell'errore  $\|x^k - x^*\|$ . In generale però questo non è vero; inoltre a volte anche spostamenti piccoli possono produrre grossi cambiamenti nel valore della funzione obiettivo. Questo è il motivo per cui si utilizza congiuntamente anche la seconda condizione.

In realtà è preferibile utilizzare la seguente modifica che consente di normalizzare rispetto ai valori assunti da  $f$  e da  $\|x\|$  (con un controllo per evitare la divisione per zero):

$$\frac{\|x^{k+1} - x^k\|}{\max\{\|x^k\|, 1\}} \leq \varepsilon_2 \quad \frac{|f(x^{k+1}) - f(x^k)|}{\max\{|f(x^k)|, 1\}} \leq \varepsilon_3 \quad (93)$$

Le condizioni (92) e (93) sono conosciute come *criterio di Himmelblau*. Valori tipici per  $\varepsilon_1, \varepsilon_2, \varepsilon_3$  sono  $10^{-3} - 10^{-5}$ .

Per quanto riguarda la scelta della norma, tra le più utilizzate ci sono  $\|\cdot\|_2$  e  $\|\cdot\|_\infty$ .

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 esistenti, una delle distinzioni più significative è quella che fa riferimento alle informazioni disponibili sulla funzione da ottimizzare ai fini del calcolo di  $d^k$ . In particolare, possiamo distinguere:

- metodi che utilizzano soltanto le derivate prime (*metodo del gradiente, metodi delle direzioni coniugate, metodi Quasi-Newton*);
- metodi che utilizzano la conoscenza delle derivate prime e delle derivate seconde (*Metodo di Newton e relative modifiche*);
- metodi senza derivate, che si basano esclusivamente sulla valutazione della funzione obiettivo lungo direzioni di ricerca prefissate (come, ad esempio, gli assi coordinati) o definite in base ai valori della funzione obiettivo nei punti precedenti.

Nel seguito considereremo solo il metodo del gradiente, il metodo del gradiente coniugato e il metodo di Newton.

4. **Calcolo del passo.** Il calcolo dello scalare  $\alpha^k$  costituisce la cosiddetta *ricerca unidimensionale* o *ricerca di linea* (*line search*) e viene effettuato valutando la funzione obiettivo (ed eventualmente le derivate prime) lungo la direzione  $d^k$ . Nel caso in cui la direzione di ricerca sia una direzione di discesa, e in particolare che soddisfi la condizione  $\nabla f(x^k)^T d^k < 0$ , potremo limitarci a considerare valori di  $\alpha > 0$ .

Possiamo facilmente dimostrare un primo risultato per lo schema di algoritmo definito sopra.

**Teorema 39** [3] *Sia  $\{x^k\}$  la successione prodotta dall' algoritmo. Supponiamo che  $\mathcal{L}_{x^0}$  sia compatto e che  $f(x^{k+1}) < f(x^k)$ . Allora:*

- (a) *la successione  $\{x^k\}$  rimane in  $\mathcal{L}_{x^0}$  ed ammette punti di accumulazione;*
- (b) *ogni punto di accumulazione di  $\{x^k\}$  appartiene a  $\mathcal{L}_{x^0}$ ;*
- (c) *la successione  $\{f(x^k)\}$  converge.*

**Dimostrazione [3]. (NON in programma)** Poiché per definizione di algoritmo di discesa  $f(x^{k+1}) < f(x^k)$  per ogni  $k$ , la successione rimane in  $\mathcal{L}_{x^0}$ . Inoltre per la compattezza di  $\mathcal{L}_{x^0}$ ,  $\{x^k\}$  ammette almeno una sottosuccessione  $\{x^k\}_K$  convergente e tutte le sottosuccessioni convergenti hanno punti di accumulazione in  $\mathcal{L}_{x^0}$ . Inoltre poiché  $f$  è continua e  $\mathcal{L}_{x^0}$  è compatto, esiste il minimo  $x^*$  di  $f$  su  $\mathcal{L}_{x^0}$  e risulta quindi  $f(x^k) \geq f(x^*)$  per ogni  $k$ . La successione  $\{f(x^k)\}$  è monotona decrescente e limitata inferiormente e quindi converge.  $\square$

Il risultato di convergenza riportato nel Teorema 39 non consente in generale di mettere in relazione i punti di accumulazione della successione  $\{x^k\}$  prodotta dall' algoritmo con i punti dell'insieme bersaglio  $\Omega$ .

In particolare la richiesta (91) di semplice riduzione della funzione  $f$ , non è sufficiente a garantire la convergenza alla soluzione, come provato nell'Esempio 15 nel paragrafo 10.3.

La scelta del passo deve garantire alcune proprietà aggiuntive. Nel seguito riportiamo due criteri per la determinazione del passo che consentono di garantire la convergenza dei metodi che considereremo successivamente.

## 10.2 Metodi di ricerca unidimensionale

Gli algoritmi di *ricerca unidimensionale* hanno per obiettivo la determinazione del passo  $\alpha^k$  lungo la direzione assegnata  $d^k$ .

In termini molto generali, si può dire che la ricerca unidimensionale è finalizzata, essenzialmente, ad assicurare la convergenza globale dell' algoritmo, mentre la rapidità di convergenza dipende in prevalenza dalla scelta della direzione.

Fissati quindi  $x^k, d^k$ , la scelta dello scalare  $\alpha$  è fatta sulla base dell'andamento di  $f$  lungo la direzione di ricerca  $d^k$ . In particolare, indicheremo con  $\phi : R \rightarrow R$  la funzione di una sola variabile  $\alpha$  definita da

$$\phi(\alpha) = f(x^k + \alpha d^k).$$

Se  $f$  è differenziabile, esiste la derivata di  $\phi(\alpha)$  rispetto ad  $\alpha$ , indicata con  $\dot{\phi}(\alpha)$ , che si può porre nella forma:

$$\dot{\phi}(\alpha) = \nabla f(x^k + \alpha d^k)^T d^k, \quad (94)$$

e coincide, ovviamente, con la derivata direzionale di  $f$  lungo la direzione  $d^k$  nel punto  $x^k + \alpha d^k$ .

Per la ricerca unidimensionale sono disponibili diversi algoritmi che possono prevedere il calcolo del punto di minimo di  $f(x^k + \alpha d^k)$  rispetto a  $\alpha$  (*ricerca esatta*) oppure l'individuazione di un intervallo di valori accettabili ai fini dell'ottenimento della convergenza per  $\alpha^k$  (*ricerca inesatta*).

### 10.2.1 Metodi di ricerca esatta

Uno dei criteri proposti per la determinazione di  $\alpha^k$  è quello di scegliere un valore di  $\alpha$  che minimizzi  $\phi(\alpha)$  lungo la direzione  $d^k$ , ossia un valore  $\alpha^k$  per cui si abbia:

$$f(x^k + \alpha^k d^k) \leq f(x^k + \alpha d^k), \quad \text{per ogni } \alpha \in R. \quad (95)$$

Se si assume che l'insieme di livello  $\mathcal{L}_{x^0}$  sia compatto e che  $x^k \in \mathcal{L}_{x^0}$ , esiste un valore di  $\alpha$  che risolve il problema (95) cioè esiste

$$\alpha^k = \arg \min_{\alpha} \phi(\alpha).$$

Inoltre, se  $\nabla f(x^k)'d^k < 0$ , ossia se  $d^k$  è una direzione di discesa, si può supporre  $\alpha \geq 0$  nella (95).

Poiché  $f$  è differenziabile, dovrà risultare:

$$\dot{\phi}(\alpha^k) = \nabla f(x^k + \alpha^k d^k)^T d^k = 0. \quad (96)$$

Dal punto di vista geometrico, la (96) esprime il fatto che se  $\alpha^k$  minimizza  $\phi(\alpha)$ , il gradiente di  $f$  in  $x^{k+1}$  deve essere ortogonale alla direzione  $d^k$ .

In generale non è facile determinare il passo  $\alpha^k$  che minimizza esattamente la funzione  $\phi(\alpha)$  ad eccezione di alcuni casi particolari.

### 10.2.2 Ricerca di linea esatta nel caso quadratico

Un caso semplice in cui è possibile determinare analiticamente la soluzione del problema (95), corrisponde alla minimizzazione di una funzione quadratica strettamente convessa

$$f(x) = \frac{1}{2}x^T Qx + c^T x \quad Q \succ 0.$$

In questo caso, possiamo scrivere uno sviluppo esatto di  $f(x^k + \alpha d^k)$  e otteniamo

$$\phi(\alpha) = f(x^k + \alpha d^k) = f(x^k) + \alpha \nabla f(x^k)^T d^k + \frac{1}{2} \alpha^2 d^{kT} Q d^k \quad (97)$$

Anche  $\phi$  è una funzione quadratica strettamente convessa della variabile  $\alpha$ , il cui minimo si trova annullando il gradiente, ovvero:

$$\dot{\phi}(\alpha) = \nabla f(x^k)^T d^k + \alpha d^{kT} Q d^k = 0$$

e quindi si ottiene:

$$\alpha^k = -\frac{\nabla f(x^k)^T d^k}{d^{kT} Q d^k} = -\frac{(Qx^k + c)^T d^k}{d^{kT} Q d^k}. \quad (98)$$

Osserviamo che nel caso di funzione quadratica dalla (97), possiamo sempre scrivere:

$$f(x^k + \alpha d^k) - f(x^k) = \alpha \nabla f(x^k)^T d^k + \frac{1}{2} \alpha^2 d^{kT} Q d^k$$

Sostituendo l'espressione del passo  $\alpha^k$  ottenuto con ricerca di linea esatta si ottiene:

$$f(x^{k+1}) - f(x^k) = -\frac{(\nabla f(x^k)^T d^k)^2}{d^{kT} Q d^k} + \frac{1}{2} \frac{(\nabla f(x^k)^T d^k)^2}{(d^{kT} Q d^k)^2} d^{kT} Q d^k = -\frac{1}{2} \frac{(\nabla f(x^k)^T d^k)^2}{d^{kT} Q d^k}$$

Quindi nel caso di minimizzazione di funzione quadratica con ricerca di linea esatta, si ottiene non solo che  $f(x^{k+1}) < f(x^k)$ , ma l'entità della riduzione della funzione lungo la direzione  $d^k$ .

**Esempio 14** Data la funzione dell'Esempio 11 del paragrafo 7.4, scrivere l'espressione di  $\phi(\alpha)$  e  $\dot{\phi}(\alpha)$  con  $x^k, d^k$  generici.

**Soluzione.** Utilizziamo l'espressione (97). Abbiamo già calcolato  $\nabla f$  e  $\nabla^2 f$ . Indichiamo  $x^k = (x_1 \ x_2)^T, d^k = (d_1 \ d_2)^T$ .

$$\phi(\alpha) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2 + \alpha \begin{pmatrix} 4x_1 + 2x_2 + 1 \\ 2x_2 + 2x_1 - 1 \end{pmatrix}^T \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} + \frac{1}{2}\alpha^2 \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}^T \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$$

e

$$\dot{\phi}(\alpha) = d_1(4x_1 + 2x_2 + 1) + d_2(2x_2 + 2x_1 - 1) + \alpha(2d_1^2 + d_2^2 + 2d_1d_2)$$

da cui si ottiene l'espressione

$$\alpha^k = -\frac{d_1(4x_1 + 2x_2 + 1) + d_2(2x_2 + 2x_1 - 1)}{2d_1^2 + d_2^2 + 2d_1d_2}$$

□

**Esercizio 11** Data la funzione dell'Esercizio 6. Calcolare l'espressione del passo  $\alpha$  ottenuto con una ricerca di linea esatta a partire dal punto  $x = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  lungo la direzione

$$d = \begin{pmatrix} 1 \\ 1/2 \end{pmatrix}.$$

### 10.2.3 Metodi di ricerca inesatta: metodo di Armijo

In pratica, la *ricerca di linea esatta*, ossia la determinazione di un  $\alpha^k$  che risolva il problema (95) è notevolmente dispendiosa, soprattutto per problemi in cui  $f$  è non convessa. D'altra parte, non esistono particolari vantaggi, dal punto di vista del funzionamento complessivo dei metodi di minimizzazione, nel determinare il punto di minimo lungo  $d^k$ . Si preferisce quindi adottare criteri "approssimati" che siano più facilmente utilizzabili dal punto di vista computazionale e garantiscano al tempo stesso opportune proprietà di convergenza.

Il *metodo di Armijo* può essere descritto per mezzo dello schema iterativo seguente:

#### Metodo di Armijo

**Dati:**  $\delta \in (0, 1), \gamma \in (0, 1/2)$ .

**Passo 1.** Si sceglie  $\Delta^k$  e si pone  $\alpha = \Delta^k$  e  $j = 0$ .

**Passo 2.** Se è soddisfatta la condizione

$$f(x^k + \alpha d^k) \leq f(x^k) + \gamma \alpha \nabla f(x^k)' d^k \quad (99)$$

si assume  $\alpha^k = \alpha$ . Stop

**Passo 3.** Si pone  $\alpha = \delta \alpha, j = j + 1$  e si ritorna al Passo 2.

Valori tipici dei parametri sono  $\delta \in [0.1, 0.5]$  e  $\gamma \in [10^{-4}, 10^{-3}]$ .

**Osservazione 9** *Si noti che nella ricerca unidimensionale è importante utilizzare una buona stima iniziale  $\Delta^k$  del passo, per ridurre il numero di valutazioni di funzione ed evitare fenomeni di overflow. Una scelta appropriata di  $\Delta^k$  deve essere tuttavia effettuata in relazione al metodo utilizzato per il calcolo di  $d^k$ .*

È immediato rendersi conto che, dal punto di vista geometrico, la condizione (99) impone che  $\alpha^k$  sia tale da assicurare che  $\phi(\alpha^k)$  sia al di sotto della retta passante per  $(0, \phi(0))$  e avente pendenza  $\gamma\dot{\phi}(0)$ , ovvero (vedi Figura 14):

$$y(\alpha) = \phi(0) + \gamma\dot{\phi}(0)\alpha.$$

La (99) si può scrivere  $\phi(\alpha) \leq y(\alpha)$  ed impone quindi un *sufficiente riduzione* della funzione obiettivo.

Una condizione di *sufficiente spostamento* viene imposta implicitamente, in quanto l'algoritmo assicura che  $\alpha^k$  sia scelto come *il più grande* valore di  $\alpha$  per cui la (99) è soddisfatta nell'insieme:

$$A = \{\alpha : \alpha = \delta^j \Delta^k, j = 0, 1, \dots\}.$$

La condizione  $\gamma < 1/2$  è motivata dall'opportunità di assicurare che, se  $f$  è una funzione quadratica, risulti accettabile il valore del passo che minimizza  $\phi(\alpha)$  ed è importante, in alcuni metodi, ai fini della rapidità di convergenza.

**Esercizio 12** *Dimostrare che se  $f$  è una funzione quadratica strettamente convessa, il passo (98) ottenuto con ricerca di linea esatta è accettabile secondo il criterio di Armijo.*

### 10.3 Il metodo del gradiente

Il *metodo del gradiente* (o *metodo della discesa più ripida*) è uno dei primi metodi proposti per la minimizzazione non vincolata e si basa sull'uso della direzione di ricerca

$$d^k = -\nabla f(x^k),$$

ossia della direzione opposta a quella del gradiente, o *antigradiente* di  $f$  in  $x^k$ .

Osserviamo che in questo caso risulta

$$\nabla f(x^k)^T d^k = -\|\nabla f(x^k)\|^2$$

e quindi se  $\nabla f(x^k) \neq 0$ , la direzione dell'antigradiente è sempre di discesa.

L'interesse della direzione  $-\nabla f(x^k)$  risiede proprio nel fatto che, se il gradiente è continuo, come si è ipotizzato, essa costituisce una direzione di discesa che si annulla se e solo se  $x$  è un punto stazionario.

Questa proprietà assicura che con una scelta opportuna del passo  $\alpha^k$  sia possibile stabilire facilmente un risultato di convergenza globale.

Il metodo del gradiente costituisce quindi il modello più significativo di algoritmo *globalmente convergente*.

Si può schematizzare la  $k$ -esima iterazione del metodo del gradiente come segue:

$$x^{k+1} = x^k - \alpha^k \nabla f(x^k), \quad (100)$$

in cui rimane da specificare come scegliere il passo  $\alpha^k$ .

Osserviamo innanzitutto come la scelta del passo  $\alpha^k$  tale da garantire una semplice riduzione della funzione obiettivo  $f(x^{k+1}) < f(x^k)$ , non sia sufficiente a garantire la convergenza del metodo.

**Esempio 15 (Non in programma)** *Sia data la funzione  $f : R \rightarrow R$*

$$f(x) = \begin{cases} \frac{3}{4}(1-x)^2 - 2(1-x) = \frac{1}{4}(3x^2 + 2x - 5) & \text{se } x > 1 \\ \frac{3}{4}(1+x)^2 - 2(1+x) = \frac{1}{4}(3x^2 - 2x - 5) & \text{se } x < -1 \\ x^2 - 1 & \text{se } -1 \leq x \leq 1 \end{cases}$$

il cui gradiente è

$$\frac{df(x)}{dx} = \begin{cases} \frac{3}{2}x + \frac{1}{2} & \text{se } x > 1 \\ \frac{3}{2}x - \frac{1}{2} & \text{se } x < -1 \\ 2x & \text{se } -1 \leq x \leq 1 \end{cases}$$

*Si tratta di una funzione continuamente differenziabile, convessa che ammette come unico punto di minimo il punto  $x^* = 0$ .*

*Mostrare che il metodo del gradiente*

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

*in cui il passo  $\alpha \leq 1$  è scelto in modo da soddisfare un semplice criterio di riduzione della funzione obiettivo non converge al punto  $x^*$ .*

È necessario quindi definire quali regole per la scelta del passo  $\alpha$  consentono di garantire la convergenza. Consideriamo dapprima il caso quadratico e poi il caso di funzione più generale.

## 10.4 Il metodo del gradiente per funzioni quadratiche

Consideriamo il caso di minimizzazione di una funzione quadratica  $f(x) = \frac{1}{2}x^T Qx + c^T x$  strettamente convessa  $Q \succ 0$ . Abbiamo visto che in questo caso è possibile calcolare analiticamente il passo  $\alpha^k$  con una ricerca di linea esatta e vale (98). Sostituendo l'espressione di  $d^k$  otteniamo il passo

$$\alpha^k = \frac{\|\nabla f(x^k)\|^2}{\nabla f(x^k)^T Q \nabla f(x^k)}$$

Riassumendo:

### Metodo del gradiente con ricerca di linea esatta

**Passo 1.** Si fissa un punto iniziale  $x^0 \in R^n$  e si pone  $k = 0$ .

**Passo 2.** Si calcola  $\nabla f(x^k)$ ; se  $\nabla f(x^k) = 0$  stop.

**Passo 3.** Si assume

$$x^{k+1} = x^k - \frac{\|\nabla f(x^k)\|^2}{\nabla f(x^k)^T Q \nabla f(x^k)} \nabla f(x^k),$$

di pone  $k = k + 1$  e si ritorna al passo 2.

Osserviamo però che il metodo del gradiente con ricerca di linea esatta si muove lungo direzioni successive che sono ortogonali.

Possiamo dimostrare il seguente risultato.

**Teorema 40** *Si consideri la funzione*

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

con  $Q$  matrice simmetrica e definita positiva. Il metodo del gradiente con ricerca unidimensionali esatta converge all'unico punto di minimo  $x^* = -Q^{-1}c$ .

**Dimostrazione.** Nel caso di funzione quadratica possiamo scrivere:

$$f(x^k + \alpha d^k) - f(x^k) = \alpha \nabla f(x^k)^T d^k + \frac{1}{2} \alpha^2 d^{kT} Q d^k$$

Sostituendo l'espressione del passo  $\alpha^k$  ottenuto con ricerca di linea esatta e della direzione  $d^k$ , si ottiene:

$$f(x^{k+1}) - f(x^k) = -\frac{1}{2} \frac{\|\nabla f(x^k)\|^4}{\nabla f(x^k)^T Q \nabla f(x^k)} < 0$$

poiché  $Q \succ 0$  e quindi il termine  $\nabla f(x^k)^T Q \nabla f(x^k) > 0$ .

Quindi la successione  $\{x^k\}$  è tale che  $f(x^{k+1}) < f(x^k)$  e dunque si mantiene in  $\mathcal{L}_{x^0}$  che è compatto. La successione  $\{x^k\}$  ammette quindi almeno una sottosuccessione convergente; inoltre poiché esiste il punto di minimo  $x^*$  e risulta  $f(x) \geq f(x^*)$  per ogni  $x$ , possiamo anche dire che la successione  $\{f(x^k)\}$  è limitata inferiormente dal valore  $f(x^*)$ . Dunque  $\{f(x^k)\}$  è una successione monotona decrescente e limitata inferiormente, e quindi, per la Proposizione 39,  $\{f(x^k)\}$  converge ad un valore  $\bar{f}$ . Si avrà quindi

$$\lim_{k \rightarrow \infty} (f(x^{k+1}) - f(x^k)) = 0.$$

Inoltre per ogni forma quadratica, vale:

$$\lambda_m \|x\|^2 \leq x^T Q x \leq \lambda_M \|x\|^2 \quad \forall x \in R^n$$

dove  $\lambda_m, \lambda_M$  indicano rispettivamente l'autovalore minimo e massimo della matrice  $Q$  e quindi in particolare possiamo anche scrivere

$$\lambda_m \|\nabla f(x^k)\|^2 \leq \nabla f(x^k)^T Q \nabla f(x^k) \leq \lambda_M \|\nabla f(x^k)\|^2 \quad \forall k$$

Ricordando che per ipotesi  $Q$  è definita positiva, si ottiene

$$f(x^{k+1}) - f(x^k) \leq -\frac{1}{2\lambda_M} \|\nabla f(x^k)\|^2.$$

Al limite otteniamo quindi

$$0 = \lim_{k \rightarrow \infty} (f(x^{k+1}) - f(x^k)) \leq -\frac{1}{2\lambda_M} \lim_{k \rightarrow \infty} \|\nabla f(x^k)\|^2$$

e quindi

$$\lim_{k \rightarrow \infty} \|\nabla f(x^k)\| = 0$$

Ovvero ogni successione  $\{x^k\}_K$  convergente, converge al punto  $x^*$ . Poiché  $x^*$  è l'unico punto stazionario, allora tutta la successione  $\{x^k\}$  converge a  $x^*$ .  $\square$

Per quanto riguarda la rapidità di convergenza del metodo del gradiente, vale la stima espressa dalla seguente proposizione.

**Teorema 41** *Si consideri la funzione*

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

con  $Q$  matrice simmetrica e definita positiva. Il metodo del gradiente con ricerca unidimensionali esatta converge all'unico punto di minimo  $x^* = -Q^{-1}c$  e si ha:

$$\|x^{k+1} - x^*\|_2 \leq \left(\frac{\lambda_M}{\lambda_m}\right)^{1/2} \left(\frac{\lambda_M - \lambda_m}{\lambda_M + \lambda_m}\right) \|x^k - x^*\|_2, \quad (101)$$

in cui  $\lambda_M$  e  $\lambda_m$  sono, rispettivamente, il massimo e il minimo autovalore di  $Q$ .

**Dimostrazione.** Usando la formula di Taylor possiamo scrivere (tenendo conto che  $\nabla f(x^*) = 0$ )

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

Sfruttando l'ipotesi di convessità otteniamo

$$\frac{\lambda_m}{2} \|x - x^*\|^2 \leq f(x) - f(x^*) \leq \frac{\lambda_M}{2} \|x - x^*\|^2. \quad (102)$$

Posto  $x(\alpha) = x^k - \alpha \nabla f(x^k)$ , possiamo scrivere

$$\begin{aligned} \|x(\alpha) - x^*\|^2 &= \left(x^k - \alpha \nabla f(x^k) - x^*\right)^T (x(\alpha) - x^*) = \\ &= \left[x^k - x^* - \alpha (\nabla f(x^k) - \nabla f(x^*))\right]^T (x(\alpha) - x^*). \end{aligned}$$

Usando la formula di Lagrange possiamo scrivere  $(\nabla f(x^k) - \nabla f(x^*))^T (x(\alpha) - x^*) = (x^k - x^*)^T Q(x(\alpha) - x^*)$ , e quindi

$$\begin{aligned} \|x(\alpha) - x^*\|^2 &= \left(x^k - x^*\right)^T (I - \alpha Q) (x(\alpha) - x^*) \\ &\leq \|I - \alpha Q\| \|x^k - x^*\| \|x(\alpha) - x^*\| \end{aligned}$$

da cui

$$\begin{aligned} \|x(\alpha) - x^*\| &\leq \|I - \alpha Q\| \|x^k - x^*\| \\ &\leq \max\{|1 - \alpha\lambda_m|, |1 - \alpha\lambda_M|\} \|x^k - x^*\| = \varphi(\alpha) \|x^k - x^*\|. \end{aligned}$$

La funzione  $\varphi(\alpha)$  assume valore minimo nei punti in cui  $1 - \alpha\lambda_m = -(1 - \alpha\lambda_M)$  ovvero per

$$\alpha_1 = \frac{2}{\lambda_M + \lambda_m}, \quad \varphi(\alpha_1) = \frac{\lambda_M - \lambda_m}{\lambda_M + \lambda_m}.$$

Quindi possiamo scrivere

$$\|x(\alpha_1) - x^*\| \leq \left( \frac{\lambda_M - \lambda_m}{\lambda_M + \lambda_m} \right) \|x^k - x^*\|. \quad (103)$$

Dalla disequazione sinistra di (102), tenuto conto che  $\alpha^k$  è ottenuto con una ricerca unidimensionale esatta e quindi  $f(x^{k+1}) \leq f(x(\alpha))$  per ogni  $\alpha$ , risulta:

$$\frac{\lambda_m}{2} \|x^{k+1} - x^*\|^2 \leq f(x^{k+1}) - f(x^*) \leq f(x(\alpha_1)) - f(x^*).$$

Utilizzando la disequazione destra della (102) e la (103) abbiamo ancora

$$\|x^{k+1} - x^*\|^2 \leq \frac{2}{m} [f(x(\alpha_1)) - f(x^*)] \leq \frac{\lambda_M}{\lambda_m} \|x(\alpha_1) - x^*\|^2 \leq \frac{\lambda_M}{\lambda_m} \left( \frac{\lambda_M - \lambda_m}{\lambda_M + \lambda_m} \right)^2 \|x^k - x^*\|^2$$

ovvero

$$\|x^{k+1} - x^*\| \leq \left( \frac{\lambda_M}{\lambda_m} \right)^{1/2} \left( \frac{\lambda_M - \lambda_m}{\lambda_M + \lambda_m} \right) \|x^k - x^*\|$$

□

La stima data nella Proposizione 41 mostra che dipende, nel caso quadratico, dal rapporto

$$r = \frac{\lambda_M}{\lambda_m} \geq 1$$

tra il massimo e il minimo autovalore della matrice Hessiana  $Q$ . Si può scrivere

$$\|x^{k+1} - x^*\|_2 \leq r^{1/2} \left( \frac{r-1}{r+1} \right) \|x^k - x^*\|_2.$$

Osserviamo inoltre che, se  $r = 1$ , la convergenza viene ottenuta in un solo passo da ogni punto iniziale. Se  $r > 1$ , la convergenza può ancora essere ottenuta in un solo passo a partire da punti situati sugli assi dell'iperellissoide che definisce una superficie di livello di  $f$ , mentre esistono punti (in particolare quelli situati in prossimità dell'asse maggiore) a partire dai quali la convergenza può essere notevolmente lenta (vedi Figura ??).

La rapidità di convergenza del metodo del gradiente peggiora, in genere, al crescere di  $r$ , ossia all'aumentare del *mal condizionamento* della matrice  $Q$ .

**Esempio 16** *Sia dato il problema di minimizzazione*

$$\min x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2.$$

*Si scrivano le prime iterazioni del metodo del metodo del gradiente con una ricerca di linea esatta a partire dal punto iniziale*

$$x^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

**Soluzione.** Si tratta della funzione studiata nell'esempio 11 il cui gradiente è

$$\nabla f(x) = \begin{pmatrix} 4x_1 + 2x_2 + 1 \\ 2x_2 + 2x_1 - 1 \end{pmatrix}$$

**Iterazione 0.**

$$d^0 = -\nabla f(x^0) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Si determina il passo

$$\alpha^0 = \arg \min_{\alpha} f(x^0 + \alpha d^0) = \arg \min_{\alpha} \phi(\alpha) = \alpha(\alpha - 2).$$

Annullando la derivata  $\dot{\phi}(\alpha) = 2\alpha - 2 = 2(\alpha - 1) = 0$ , si ottiene  $\alpha^0 = 1$  e quindi

$$x^1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

**Iterazione 1.** Verifica condizioni di arresto

$$\nabla f(x^1) = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \neq 0$$

quindi il punto non è ottimo. La nuova direzione è

$$d^1 = -\nabla f(x^1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Determiniamo  $\alpha^1 = \arg \min_{\alpha} f(x^1 + \alpha d^1) = \arg \min_{\alpha} (5\alpha^2 - 2\alpha - 1)$ . Annullando la derivata

$$\dot{\phi}(\alpha) = 10\alpha - 2 = 0$$

si ottiene  $\alpha^1 = \frac{1}{5}$ . Quindi

$$\begin{aligned} x^2 &= x^1 + \alpha^1 \nabla f(x^1) = \\ &= \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1/5 \\ 1/5 \end{bmatrix} = \begin{bmatrix} -0.8 \\ 1.2 \end{bmatrix} \end{aligned}$$

**Iterazione 2.** Verifica condizione di arresto

$$\nabla f(x^2) = \begin{bmatrix} 1/5 \\ -1/5 \end{bmatrix} \neq 0,$$

quindi il punto  $x^2$  non è ottimo. Si calcola la direzione

$$d^2 = -\nabla f(x^2) = \begin{bmatrix} -\frac{1}{5} \\ \frac{1}{5} \end{bmatrix} = \begin{pmatrix} -0.2 \\ 0.2 \end{pmatrix}.$$

Determiniamo  $\alpha^2 = \arg \min_{\alpha} f(x^2 + \alpha d^2) = \arg \min_{\alpha} (\alpha^2 - 2\alpha - 30)/25$ . Annullando la derivata

$$\dot{\phi}(\alpha) = 2\alpha - 2 = 0$$

si ottiene  $\alpha^2 = 1$ . Quindi il punto  $x^3 = x^2 + \alpha^2 d^2$  è

$$x^3 = \begin{bmatrix} -1 \\ 7/5 \end{bmatrix} = \begin{bmatrix} -1 \\ 1,4 \end{bmatrix}.$$

**Iterazione 3.** Verifica condizione di arresto:

$$\nabla q(x^3) = \begin{bmatrix} -\frac{1}{5} \\ -\frac{1}{5} \end{bmatrix} \neq 0$$

e il punto non è ottimo. Si dovrebbe proseguire.

Noi terminiamo qui l'applicazione dell'algoritmo. La rapidità di convergenza è determinata dal rapporto

$$\frac{\lambda_{\max}(Q)}{\lambda_{\min}(Q)} = \left( \frac{3 + \sqrt{5}}{3 - \sqrt{5}} \right).$$

□

**Esercizio 13** Si consideri la funzione dell'esercizio 6. Si applichi il metodo del gradiente con ricerca di linea esatta a partire dal punto  $x^0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ .

Nel caso di funzione quadratica si può garantire la convergenza del metodo del gradiente in un numero finito di iterazioni scegliendo il passo  $\alpha^k$  in modo opportuno.

In particolare possiamo definire il metodo (**NON in PROGRAMMA**)

#### Metodo del gradiente con uso degli autovalori

**Passo 1.** Si fissa un punto iniziale  $x^0 \in R^n$ , si calcolano gli autovalori di  $Q$

$$\lambda_1 \leq \dots \leq \lambda_n.$$

**Do k=0, ..., n-1**

**Passo k2.** Si calcola  $\nabla f(x^k)$ ; se  $\nabla f(x^k) = 0$  stop.

**Passo k3.** Si pone

$$x^{k+1} = x^k - \frac{1}{\lambda_{k+1}} \nabla f(x^k), \quad (104)$$

**End Do**

Vale il seguente risultato che si riporta senza dimostrazione.

**Teorema 42** Si consideri la funzione

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

con  $Q$  matrice simmetrica e definita positiva e siano  $\lambda_1 \leq \dots \leq \lambda_n$  gli autovalori di  $Q$ . Il metodo del gradiente, descritto alla generica iterazione  $k \geq 0$  dalla formula di aggiornamento (104), converge al massimo in  $n$  iterazioni all'unico punto stazionario  $x^* = -Q^{-1}c$  di  $f$ .

**Esempio 17** Sia dato il problema dell'Esercizio 16. Si scrivano le iterazioni del metodo del metodo del gradiente del tipo (104) a partire dal punto iniziale  $x^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ .

**Soluzione.** Abbiamo già calcolato gli autovalori di  $Q$  che sono  $\lambda_1 = 3 - \sqrt{5}$  e  $\lambda_2 = 3 + \sqrt{5}$ . Quindi le iterazioni sono

Iterazione 1.

$$x^1 = x^0 - \frac{1}{\lambda_1} \nabla f(x^0)$$

$$x^1 = x^0 - \frac{1}{\lambda_1} \nabla f(x^0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \frac{1}{3 - \sqrt{5}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{3 + \sqrt{5}}{4} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Iterazione 2.

$$x^2 = x^1 - \frac{1}{\lambda_2} \nabla f(x^1) = \frac{3 + \sqrt{5}}{4} \begin{pmatrix} -1 \\ 1 \end{pmatrix} - \frac{1}{3 + \sqrt{5}} \begin{pmatrix} -\frac{\sqrt{5}+1}{2} \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 3/2 \end{pmatrix}$$

Il minimo è stato ottenuto in  $n = 2$  iterazioni.

Osserviamo che possiamo scrivere:

$$x^2 = x^1 - \frac{1}{\lambda_2} \nabla f(x^1) = x^0 - \frac{1}{\lambda_1} \nabla f(x^0) - \frac{1}{\lambda_2} \nabla f(x^1)$$

Generalizzando al caso  $n$  dimensionale, per una funzione quadratica strettamente convessa possiamo scrivere analiticamente, l'espressione della soluzione ottima che è:

$$x^* = x^0 - \sum_{j=0}^{n-1} \frac{1}{\lambda_{j+1}} \nabla f(x^j).$$

□

Quindi utilizzando “maggior” informazioni, in particolare informazioni del secondo ordine (gli autovalori della matrice  $Q$ ), si riesce ad ottenere una maggiore rapidità di convergenza finita. Vedremo nel paragrafo dedicato al metodo del gradiente coniugato che questo stesso tipo di convergenza si può ottenere utilizzando implicitamente informazioni del secondo ordine nella definizione della direzione di ricerca piuttosto che nella definizione del passo.

**Esercizio 14** Si consideri la funzione dell'Esercizio 6. Mostrare che il metodo del gradiente definito dall'iterazione (104) produce le stesse iterazioni del metodo del gradiente con ricerca di linea esatta. Dare una giustificazione teorica.

Consideriamo ora il caso di funzioni non quadratiche.

## 10.5 Il metodo del gradiente con ricerca di linea

Nel caso di funzione non quadratica, abbiamo già osservato che effettuare una ricerca di linea esatta può essere troppo oneroso e non necessario.

Per il metodo del gradiente si può utilizzare una ricerca di linea di tipo Armijo come descritto nel paragrafo 10.2.3. In particolare definiamo il seguente schema generale di algoritmo del metodo del gradiente.

### Metodo del gradiente con ricerca di linea di Armijo

**Dati.**  $\gamma \in (0, 1/2)$  e  $\delta \in (0, 1)$ .

**Passo 1.** Si fissa un punto iniziale  $x^0 \in R^n$  e si pone  $k = 0$ .

**Passo 2.** Si calcola  $\nabla f(x^k)$ ; se  $\nabla f(x^k) = 0$  stop.

**Passo 3.** Si sceglie  $\alpha = \Delta^k$  e si calcola il più piccolo valore intero  $j^* \in \{0, 1, 2, \dots\}$  tale che

$$f(x^{k+1}) \leq f(x^k) - \gamma \alpha \delta^j \|\nabla f(x^k)\|^2,$$

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

Si pone  $\alpha^k = \alpha \delta^{j^*}$ .

**Passo 4.** Si assume

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

$k = k + 1$  e si ritorna al passo 2.

Vale la seguente proposizione che assicura la convergenza globale del metodo del gradiente con una tecnica di ricerca unidimensionale tipo-Armijo.

**Teorema 43** [3] *Supponiamo che valga l'Ipotesi 1. Sia  $\{x^k\}$  la successione prodotta dal metodo del gradiente con una ricerca unidimensionale di tipo-Armijo.*

*Allora, o esiste un indice  $\nu \geq 0$  tale che  $x^\nu \in \mathcal{L}_{x^0}$  e  $\nabla f(x^\nu) = 0$ , oppure viene prodotta una successione infinita tale che ogni punto di accumulazione  $\bar{x}$  di  $\{x^k\}$  appartiene a  $\mathcal{L}_{x^0}$  e soddisfa  $\nabla f(\bar{x}) = 0$ .*

Naturalmente le stesse conclusioni della Proposizione 43, valgono anche se il passo  $\alpha$  è determinato con una ricerca di linea esatta.

Per quanto riguarda la rapidità di convergenza del metodo del gradiente nel caso di funzioni non quadratiche, si può ottenere una stima analoga a quella del Teorema 40 con riferimento al caso di funzione convessa e di ricerca di linea esatta.

**Esempio 18** *Si consideri il problema dell'Esempio 11 del paragrafo 7.4. Si determini l'intervallo di valori di  $\alpha$  accettabili secondo il criterio di Armijo alla prima iterazione del metodo del gradiente con punto iniziale  $x^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ , posto  $\gamma = 1/4$ .*

**Soluzione.** Abbiamo già calcolato il gradiente nell'esempio 14. La prima iterazione del metodo del gradiente parametrica nel passo  $\alpha$  è

$$x^1(\alpha) = x^0 - \alpha \nabla f(x^0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \alpha \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -\alpha \\ \alpha \end{pmatrix}$$

La regola di Armijo richiede che risulti

$$f(x^1(\alpha)) \leq f(x^0) - \gamma \alpha \|\nabla f(x^0)\|^2$$

Sostituendo i valori otteniamo la condizione:

$$-2\alpha + \alpha^2 \leq 0 - \frac{1}{4}\alpha^2 \quad \implies \quad \alpha^2 - \frac{3}{2}\alpha \leq 0$$

he è soddisfatto per valori di  $\alpha \in (0, \frac{3}{2}]$ . □

**Esercizio 15** *Si consideri la funzione dell'Esercizio 6. Sia  $x^0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ , si determini l'intervallo di valori di  $\alpha$  accettabili secondo la ricerca di linea di Armijo alla prima iterazione del metodo del gradiente.*

## 10.6 Il metodo del gradiente coniugato

I *metodi delle direzioni coniugate* sono stati originariamente introdotti come metodi iterativi per la minimizzazione di funzioni quadratiche e sono stati successivamente estesi alla minimizzazione di funzioni non quadratiche.

Una delle caratteristiche principali dei metodi delle direzioni coniugate è che essi consentono di determinare in un numero finito di iterazioni il minimo di una funzione quadratica.

Abbiamo visto nel paragrafo precedente che un risultato analogo si può ottenere con il metodo del gradiente con una scelta opportuna del passo  $\alpha$  che richiede il calcolo degli autovalori della matrice  $Q$ .

La motivazione alla base dell'uso dei metodi delle direzioni coniugate nell'ambito dell'ottimizzazione non vincolata sta nel fatto che tali metodi consentono di ottenere una rapidità di convergenza superiore rispetto a quella del metodo del gradiente, senza tuttavia richiedere *esplicitamente* la conoscenza della matrice Hessiana e senza far uso di operazioni matriciali complesse. Ciò li rende particolarmente vantaggiosi nella soluzione di problemi di ottimizzazione a grandi dimensioni.

Nel seguito analizzeremo dapprima i metodi delle direzioni coniugate con riferimento alla minimizzazione di funzioni quadratiche; successivamente considereremo l'estensione al caso non quadratico.

## 10.7 Il metodo delle direzioni coniugate per funzioni quadratiche

Consideriamo una funzione quadratica del tipo:

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

in cui  $Q$  è una matrice simmetrica definita positiva e  $c \in R^n$ , il cui unico punto di minimo è  $x^* = -Q^{-1}c$ .

Introduciamo la seguente definizione.

**Definizione 8** *Assegnata una matrice  $Q$  simmetrica e definita positiva, due vettori non nulli  $d^i, d^j \in R^n$  si dicono coniugati rispetto a  $Q$  (oppure  $Q$ -coniugati, oppure  $Q$ -ortogonali) se risulta:*

$$d^{iT} Q d^j = 0.$$

Si può dimostrare facilmente la seguente proprietà di vettori mutuamente coniugati.

**Teorema 44** [3] *Siano  $d^0, d^1, \dots, d^m \in R^n$   $m+1$  vettori non nulli e mutuamente coniugati rispetto a una matrice  $Q$  simmetrica definita positiva  $n \times n$ . Allora  $d^0, d^1, \dots, d^m$  sono linearmente indipendenti.*

**Dimostrazione.** Siano  $\alpha_0, \alpha_1, \dots, \alpha_m$  costanti reali tali che

$$\sum_{j=0}^m \alpha_j d_j = 0. \quad (105)$$

Moltiplicando a sinistra per  $Q$  ed eseguendo il prodotto scalare per  $d_i$  si ottiene, essendo  $d_i' Q d_j = 0$  per  $i \neq j$ :

$$0 = \sum_{j=0}^m \alpha_j d_i' Q d_j = \alpha_i d_i' Q d_i.$$

D'altra parte, poichè  $Q$  è definita positiva e  $d_i \neq 0$  si ha  $d_i' Q d_i > 0$  e quindi deve essere necessariamente  $\alpha_i = 0$ . Ripetendo lo stesso ragionamento per  $i = 0, 1, \dots, m$  si può affermare che la (105) implica  $\alpha_i = 0$  per  $i = 0, 1, \dots, m$ , il che prova l'indipendenza lineare dei vettori  $d_0, \dots, d_m$ .  $\square$

Definiamo allora il metodo delle direzioni coniugate per la minimizzazione di una funzione quadratica strettamente convessa.

#### Metodo delle direzioni coniugate per funzioni quadratiche

**Passo 1.** Si fissa un punto iniziale  $x^0 \in R^n$ , si calcolano  $n$  vettori non nulli  $d^0, d^1, \dots, d^{n-1} \in R^n$  mutuamente coniugati rispetto a  $Q$ .

**Do  $k=0, \dots, n-1$**

**Passo k2.** Si calcola  $\nabla f(x^k)$ ; se  $\nabla f(x^k) = 0$  stop.

**Passo k3.** Si pone

$$x^{k+1} = x^k + \alpha^k d^k,$$

con  $\alpha^k$  ottenuto con una ricerca di linea esatta:

$$\alpha^k = -\frac{\nabla f(x^k)^T d^k}{d^{kT} Q d^k}.$$

**End Do**

Possiamo ora dimostrare la convergenza del metodo in un numero finito di passi.

**Teorema 45** *Sia  $f$  una funzione quadratica del tipo  $f(x) = \frac{1}{2}x^T Q x + c^T x$ , con  $Q$  una matrice simmetrica definita positiva e siano*

$$d^0, d^1, \dots, d^{n-1}$$

*$n$  vettori non nulli e mutuamente coniugati rispetto a  $Q$ .*

*Sia  $\{x^k\}$  la successione generata dal metodo delle direzioni coniugate. Allora esiste un intero  $m \leq n$  tale che  $x^m$  coincide con il punto di minimo  $x^*$  di  $f$ .*

**Dimostrazione.** È sufficiente dimostrare che esiste  $m \leq n$  tale che  $\nabla f(x^m) = 0$ ; infatti, per la stretta convessità di  $f$ , questo implica che  $x^m$  è il punto di minimo di  $f$ .

Sia  $k$  un intero e supponiamo che risulti (altrimenti l'algoritmo si sarebbe fermato)

$$\nabla f(x^i) \neq 0, \text{ per } i = 0, 1, \dots, k-1.$$

Sia  $x^i$  un qualunque punto iniziale, l'iterazione  $x^k$  ottenuta a partire da  $x^i$  (con  $k > i$ ) avrà la seguente espressione:

$$x^k = x^i + \sum_{j=i}^{k-1} \alpha^j d^j.$$

Premoltiplicando per  $Q$  si ottiene:

$$Qx^k = Qx^i + \sum_{j=i}^{k-1} \alpha^j Qd^j.$$

Poiché  $\nabla f(x) = Qx + c$ , dall'equazione precedente, aggiungendo a destra e sinistra  $c$ , si ha:

$$\nabla f(x^k) = \nabla f(x^i) + \sum_{j=i}^{k-1} \alpha^j Qd^j,$$

Moltiplicando ancora scalarmente per  $d^i$  si ha:

$$d^{iT} \nabla f(x^k) = d^{iT} \nabla f(x^i) + \sum_{j=i}^{k-1} \alpha^j d^{iT} Qd^j$$

Tenendo conto dell'ipotesi di coniugatezza delle direzioni per cui  $d^{iT} Qd^j = 0$  per  $i \neq j$ , l'unico termine della sommatoria diverso da zero è quello corrispondente a  $j = i$  e quindi si ha:

$$d^{iT} \nabla f(x^k) = d^{iT} \nabla f(x^i) + \alpha^i d^{iT} Qd^i.$$

Ricordando ora l'espressione di  $\alpha_i$  si ottiene

$$d^{iT} \nabla f(x^k) = d^{iT} \nabla f(x^i) + \alpha^i d^{iT} Qd^i = d^{iT} \nabla f(x^i) - \frac{d^{iT} \nabla f(x^i)}{d^{iT} Qd^i} d^{iT} Qd^i = 0.$$

Poiché  $x^i$  è stato scelto in modo del tutto arbitrario, purché  $k > i$  e  $\nabla f(x^i) \neq 0$ , si può scrivere:

$$\nabla f(x^k)^T d^i = 0 \quad \text{per ogni } i = 0, 1, \dots, k-1, \quad (106)$$

cioè il gradiente al passo  $k$  è ortogonale a tutte le direzioni coniugate usate fino al passo  $k$ .

Supponiamo allora che sia  $\nabla f(x^i) \neq 0$ , per  $i = 0, 1, \dots, n-1$ . Allora possiamo scrivere la relazione (106) per  $k = n$  si ha

$$\nabla f(x^n)^T d^i = 0, \quad \text{per ogni } i = 0, 1, \dots, n-1$$

e quindi  $\nabla f(x^n)$  è ortogonale agli  $n$  vettori  $d^0, \dots, d^{n-1}$  che, per la Proposizione 44 sono linearmente indipendenti, ciò implica che

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

□

**Esempio 19** Consideriamo nuovamente il problema dell'Esempio (16)

$$\min x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$$

Sia assegnato il vettore

$$d^0 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

Determinare un vettore  $d^1$  mutuamente coniugato con  $d^0$  rispetto alla matrice hessiana

$$Q = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}.$$

Applicare il metodo delle direzioni coniugate con direzioni assegnate  $d^0, d^1$ , a partire dal punto iniziale  $x^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ .

**Soluzione.** Dato il vettore  $d^0 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ , generiamo un vettore  $d^1$  coniugato cioè tale che  $d^{0T}Qd^1 = 0$ . Si ha:

$$(1 \quad -2) \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} d_1^1 \\ d_2^1 \end{pmatrix} = (0 \quad -2) \begin{pmatrix} d_1^1 \\ d_2^1 \end{pmatrix} = -2d_2^1 = 0.$$

Quindi un qualunque vettore della forma

$$d^1 = \begin{pmatrix} d_1^1 \\ 0 \end{pmatrix}$$

è  $Q$  coniugato con  $d^0$ . Scegliamo, ad esempio,

$$d^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Applichiamo ora il metodo delle direzioni coniugate con direzioni assegnate  $d^0, d^1$ .

**Iterazione 1.** Calcoliamo il passo  $\alpha^0$  che minimizza esattamente  $f(x)$  lungo  $d^0$ .

$$\alpha^0 = -\frac{\nabla f(x^0)^T d^0}{d^{0T}Qd^0} = \frac{(1 \quad -1) \begin{pmatrix} 1 \\ -2 \end{pmatrix}}{(1 \quad -2) \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \end{pmatrix}} = -\frac{3}{4}$$

Quindi il nuovo punto  $x^1 = x^0 + \alpha^0 d^0$  è

$$x^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \frac{3}{4} \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -3/4 \\ 3/2 \end{pmatrix}$$

**Iterazione 2.** Verifica condizione di arresto:

$$\nabla f(x^1) = \begin{pmatrix} 1 \\ 1/2 \end{pmatrix} \neq 0$$

quindi il punto non é ottimo e si prosegue. Minimizzando in modo esatto lungo  $d^{(1)}$  si ha:

$$\alpha^1 = -\frac{\nabla f(x^1)^T d^1}{d^{1T} Q d^1} = \frac{-(1 \quad -5/2) \begin{pmatrix} 1 \\ 0 \end{pmatrix}}{(1 \quad 0) \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}} = -\frac{1}{4}$$

Il nuovo punto é quindi Determiniamo il punto  $x^2 = x^1 + \alpha^1 d^1$  e si ottiene

$$x^2 = \begin{pmatrix} -1 \\ 3/2 \end{pmatrix}$$

che é il punto ottimo. Infatti risulta

$$\nabla f(x^2) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

La soluzione ottima é ottenuta in  $n = 2$  passi.

**Osservazione 10** Notiamo che risulta  $\alpha^0, \alpha^1 < 0$ . Infatti abbiamo

$$\nabla f(x^0)^T d^0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}^T \begin{pmatrix} 1 \\ -2 \end{pmatrix} = 3 \quad \nabla f(x^1)^T d^1 = \begin{pmatrix} 1 \\ 1/2 \end{pmatrix}^T \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1$$

cioè le direzioni  $d^0$  e  $d^1$  sono di salita.

□

**Esercizio 16** Consideriamo il problema dell'Esempio (16). Applicare il metodo delle direzioni coniugate con direzioni assegnate  $d^0 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ ,  $d^1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ , a partire dal punto iniziale  $x^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ .

**Esercizio 17** Si consideri la funzione dell'Esercizio 6. Si applichi il metodo delle direzioni coniugate con ricerca di linea esatta a partire dal punto  $x^0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  e con una direzione assegnata  $d^0 = \begin{pmatrix} 1 \\ 1/2 \end{pmatrix}$ .

## 10.8 Metodo del gradiente coniugato per funzioni quadratiche

Nella Proposizione 45 si è supposto che le direzioni  $d^0, d^1, \dots, d^{n-1}$  fossero direzioni coniugate assegnate. Il calcolo a priori di vettori mutuamente coniugati rispetto ad una matrice  $Q$  può essere però troppo oneroso e coinvolge operazioni complesse sulla matrice  $Q$ . In generale, quindi è interessante uno schema di algoritmo in cui i vettori  $Q$ -coniugati vengono generati iterativamente.

Questa classe di metodi è nota come *metodo del gradiente coniugato*. La prima direzione è quella dell'antigradiente  $d^0 = -\nabla f(x^0)$  e le direzioni successive, tutte coniugate con  $d^0$  e tra di loro, sono ottenute con una combinazione lineare della direzione

dell'antigradiente e della direzione ottenuta al passo precedente ovvero sono generate con la formula di aggiornamento

$$d^{k+1} = -\nabla f(x^{k+1}) + \beta^k d^k; \quad (107)$$

la scelta di  $\beta^k$  assicura che la direzione  $d^{k+1}$  sia coniugata rispetto a  $d^k$ , ovvero si impone che  $d^{kT} Q d^{k+1} = -d^{kT} Q \nabla f(x^{k+1}) + \beta^k d^{kT} Q d^k = 0$  da cui si ottiene l'espressione del  $\beta^k$

$$\beta^k = \frac{\nabla f(x^{k+1})^T Q d^k}{d^{kT} Q d^k}. \quad (108)$$

È possibile eliminare la presenza esplicita della matrice  $Q$  nella definizione di  $\beta^k$ ; a partire dalla formula di aggiornamento  $x^{k+1} = x^k + \alpha^k d^k$ , moltiplicando per  $Q$  e aggiungendo a destra e sinistra  $c$ , si ottiene  $Qx^{k+1} + c = Qx^k + c + \alpha^k Qd^k$  ovvero, ricordando l'espressione del gradiente,

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

da cui si può ricavare

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

Sostituendo nell'espressione di  $\beta^k$  e semplificando  $\alpha^k$ , si ottiene:

$$\beta^k = \frac{\nabla f(x^{k+1})^T (\nabla f(x^{k+1}) - \nabla f(x^k))}{d^{kT} (\nabla f(x^{k+1}) - \nabla f(x^k))} \quad (109)$$

Osserviamo che (108) e (109) sono equivalenti per funzioni quadratiche e indipendenti dalla definizione della direzione e del passo.

A partire da (109), si possono dare ulteriori diverse espressioni per il termine  $\beta^k$ . Ad esempio nel caso di uso di ricerca di linea esatta, vale la condizione  $d^{kT} \nabla f(x^{k+1}) = 0$ . Utilizzandola per esempio nella (109) si può scrivere:

$$\beta^k = -\frac{\nabla f(x^{k+1})^T (\nabla f(x^{k+1}) - \nabla f(x^k))}{d^{kT} \nabla f(x^k)}. \quad (110)$$

Inoltre dalla (107) moltiplicata per  $\nabla f(x^{k+1})$  si ottiene (per qualunque  $k$ )

$$\nabla f(x^{k+1})^T d^{k+1} = -\|\nabla f(x^{k+1})\|^2 + \beta^k \nabla f(x^{k+1})^T d^k = -\|\nabla f(x^{k+1})\|^2.$$

Sostituendo nella (110) si ottiene una formula molto famosa detta di *Polak-Ribière*:

$$\beta^k = \frac{\nabla f(x^{k+1})^T (\nabla f(x^{k+1}) - \nabla f(x^k))}{\|\nabla f(x^k)\|^2}. \quad (111)$$

Sfruttando ulteriori proprietà nel caso quadratico, che non riportiamo, si ottiene la formula di *Fletcher-Reeves*

$$\beta^k = \frac{\|\nabla f(x^{k+1})\|^2}{\|\nabla f(x^k)\|^2}. \quad (112)$$

Riportiamo un possibile schema di metodo del gradiente coniugato.

### Metodo del gradiente coniugato di Fletcher-Reeves

**Passo 1.** Si fissa un punto iniziale  $x^0 \in R^n$  e si pone

$$d^0 = -\nabla f(x^0) = -(Qx^0 + c) \quad (113)$$

**Do k=0, ..., n-1**

**Passo k2.** Si calcola  $\nabla f(x^k)$ ; se  $\nabla f(x^k) = 0$  stop.

**Passo k3.** Si pone

$$x^{k+1} = x^k + \alpha^k d^k,$$

con  $\alpha^k$  ottenuto con una ricerca di linea esatta:

$$\alpha^k = -\frac{\nabla f(x^k)^T d^k}{d^{kT} Q d^k}.$$

**Passo k4.** Se  $\nabla f(x^{k+1}) \neq 0$ , definiamo  $d^{k+1}$ :

$$d^{k+1} = -\nabla f(x^{k+1}) + \beta^k d^k,$$

in cui:

$$\beta^k = \frac{\|\nabla f(x^{k+1})\|^2}{\|\nabla f(x^k)\|^2}.$$

**End Do**

Si può dimostrare che l'algoritmo del gradiente coniugato genera direzioni coniugate e quindi determina in un numero finito di passi il punto di minimo della funzione quadratica  $f$ . In particolare vale il seguente risultato che si riporta senza dimostrazione.

**Teorema 46** [3] *Sia  $f$  una funzione quadratica del tipo  $f(x) = \frac{1}{2}x^T Qx + c^T x$ , con  $Q$  una matrice simmetrica definita positiva e sia  $\{x^k\}$  la successione generata dal metodo del gradiente coniugato. Allora le direzioni generate dal metodo del gradiente coniugato sono coniugate tra loro ed esiste un indice  $m \leq n$  tale che  $x^m$  coincide con il punto di minimo  $x^*$  di  $f$ .*

**Esempio 20** *Consideriamo il problema dell'esercizio 16*

$$\min x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2 =$$

*Applicare il metodo del gradiente coniugato a partire dal punto iniziale  $x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ .*

**Soluzione.** La prima iterazione è la stessa del metodo del gradiente vista nell'esempio 16. **Iterazione 1.** La direzione è

$$d^0 = -\nabla f(x^0) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Con una ricerca di linea esatta si ottiene quindi

$$x^1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

**Iterazione 2.** Verifica condizioni di arresto:

$$\nabla q(x^1) = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \neq 0$$

quindi si prosegue.

Si calcola una nuova direzione mutuamente coniugata con  $d^0$ ,

$$d^1 = -\nabla f(x^1) + \beta^0 d^0$$

$$\beta^0 = \frac{\|\nabla f(x^1)\|^2}{\|\nabla f(x^0)\|^2} = 1$$

$$d^{(1)} = -\begin{bmatrix} -1 \\ -1 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

Calcoliamo il passo  $\alpha^{(1)}$  secondo la formula esatta

$$\alpha^{(1)} = \frac{(1 \quad 1) \begin{pmatrix} 0 \\ 2 \end{pmatrix}}{(0 \quad 2) \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \end{pmatrix}} = \frac{1}{4}$$

che equivale a minimizzare la funzione  $\phi(\alpha) = 4\alpha^2 - 2\alpha - 1$

Quindi si ottiene il nuovo punto  $x^2 = x^1 + \alpha^1 d^1$  come

$$x^2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \frac{1}{4} \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 3/2 \end{bmatrix} = x^*$$

La soluzione ottima è stata trovata in  $n = 2$  passi.

## 10.9 Il metodo del gradiente coniugato nel caso non quadratico

Nel paragrafo precedente abbiamo ottenuto diverse formule equivalenti per lo scalare  $\beta^k$ . Tutte le formule sono state ottenute sfruttando alcune proprietà delle funzioni quadratiche e il fatto che la ricerca di linea fosse esatta. Nel caso di funzioni non quadratiche le proprietà utilizzate per derivare le diverse formule non sono più valide e, come conseguenza, le formule di  $\beta^k$  (108)(110)(112) non sono equivalenti tra loro, ma danno luogo a metodi diversi.

In particolare nelle (110) e (112) non compare esplicitamente la matrice hessiana  $Q$  e quindi possono essere utilizzate anche nel caso di funzioni non quadratiche.

Scegliamo di descrivere brevemente solo un possibile algoritmo basato sulla formula di Fletcher-Reeves.

### Metodo del gradiente coniugato di Fletcher-Reeves

**Passo 0.** Si fissa un punto iniziale  $x^0 \in R^n$  tale che  $\nabla f(x^0) \neq 0$ ; si pone

$$d^0 = -\nabla f(x^0)$$

e  $k = 0$ .

**Passo 2.** Si calcola  $\alpha^k$  con una ricerca unidimensionale e si pone

$$x^{k+1} = x^k + \alpha^k d^k.$$

**Passo 3.** Se  $\nabla f(x^{k+1}) \neq 0$ , si pone  $d^{k+1}$ :

$$d^{k+1} = -\nabla f(x^{k+1}) + \beta^k d^k,$$

in cui:

$$\beta^k = \frac{\|\nabla f(x^{k+1})\|^2}{\|\nabla f(x^k)\|^2},$$

e si torna al passo 2.

Per definire completamente lo schema precedente è necessario specificare come deve essere fatta la ricerca unidimensionale per scelta di  $\alpha^k$ . Osserviamo che nel punto  $x^{k+1}$  la direzione generata dal metodo potrebbe non essere di discesa. Infatti vale

$$\nabla f(x^{k+1})^T d^{k+1} = -\|\nabla f(x^{k+1})\|^2 + \beta^k \nabla f(x^{k+1})^T d^k.$$

Se la ricerca di linea è esatta, risulta  $\nabla f(x^{k+1})^T d^k = 0$  e quindi la direzione  $d^{k+1}$  è effettivamente di discesa poiché risulta:

$$\nabla f(x^{k+1})^T d^{k+1} = -\|\nabla f(x^{k+1})\|^2 < 0$$

Nel caso di funzioni non lineari, però effettuare una ricerca di linea esatta non è computazionalmente accettabile e quindi in generale si utilizzano criteri di ricerca di linea inesatti per cui risulta  $\nabla f(x^{k+1})^T d^k \neq 0$ . La scelta del passo  $\alpha^k$  deve quindi garantire almeno che  $\nabla f(x^{k+1})^T d^{k+1} < 0$  cioè che

$$-\|\nabla f(x^{k+1})\|^2 + \beta^k \nabla f(x^{k+1})^T d^k < 0$$

che può essere assicurato rendendo sufficientemente piccolo il termine  $\nabla f(x^{k+1})^T d^k$ , cosicché il termine dominante sia  $-\|\nabla f(x^{k+1})\|^2$ . Questo condizione NON può essere assicurata usando solo una ricerca di linea di Armijo, ma richiede una ricerca di linea più elaborata (metodo di Wolfe [3]) che sarà oggetto di successivi corsi.

### 10.10 Il metodo di Newton

Sia  $f : R^n \rightarrow R$  una funzione due volte continuamente differenziabile; il *metodo di Newton* per la minimizzazione di  $f$  consiste nel costruire una successione di punti minimizzando a ogni passo un'approssimazione quadratica della funzione.

Se  $x^k$  è un punto assegnato, si può scrivere:

$$f(x^k + s) = f(x^k) + \nabla f(x^k)'s + \frac{1}{2}s'\nabla^2 f(x^k)s + \beta(x^k, s),$$

in cui  $\beta(x^k, s)/\|s\|^2 \rightarrow 0$  per  $s \rightarrow 0$ . Per valori sufficientemente piccoli di  $\|s\|$  si può allora pensare di approssimare  $f(x^k + s)$  con la funzione quadratica

$$q^k(s) = f(x^k) + \nabla f(x^k)'s + \frac{1}{2}s'\nabla^2 f(x^k)s$$

e determinare il punto successivo

$$x^{k+1} = x^k + s^k$$

scegliendo  $s^k$  in modo da minimizzare (ove possibile) la funzione  $q^k(s)$  rispetto a  $s$ . Poiché

$$\nabla q^k(s) = \nabla f(x^k) + \nabla^2 f(x^k)s,$$

se si suppone che  $\nabla^2 f(x^k)$  sia definita positiva, il punto di minimo di  $q^k(s)$  sarà dato da:

$$s^k = -[\nabla^2 f(x^k)]^{-1}\nabla f(x^k).$$

Il metodo di Newton è allora definito dall'iterazione

$$x^{k+1} = x^k - [\nabla^2 f(x^k)]^{-1}\nabla f(x^k). \quad (114)$$

Le proprietà di convergenza locale del metodo di Newton sono state oggetto di studi approfonditi. In quel che segue, tuttavia, ci limiteremo a caratterizzare le proprietà di convergenza locale in  $R^n$  assumendo come ipotesi l'esistenza di soluzioni.

Riportiamo, in particolare, il risultato seguente (per la dimostrazione si veda [3]).

**Teorema 47** *Sia  $f : R^n \rightarrow R$  una funzione due volte continuamente differenziabile su un insieme aperto  $\mathcal{D} \subseteq R^n$ .*

*Supponiamo inoltre che valgano le condizioni seguenti:*

- (i) *esiste un  $x^* \in \mathcal{D}$  tale che  $\nabla f(x^*) = 0$ ;*
- (ii) *la matrice hessiana  $\nabla^2 f(x^*)$  è non singolare;*
- (iii) *esiste una costante  $L > 0$  tale che, per ogni  $x, y \in \mathcal{D}$ , si abbia*

$$\|\nabla^2 f(x) - \nabla^2 f(y)\| \leq L\|x - y\|.$$

*Allora esiste una sfera aperta  $\mathcal{B}(x^*; \varepsilon) \subset \mathcal{D}$ , tale che, se  $x^0 \in \mathcal{B}(x^*; \varepsilon)$ , la successione  $\{x^k\}$  generata dal metodo di Newton (114) rimane in  $\mathcal{B}(x^*; \varepsilon)$  e converge a  $x^*$  con rapidità di convergenza quadratica.*

Si può osservare che il risultato precedente caratterizza la convergenza locale del metodo di Newton nell'intorno di un qualsiasi *punto stazionario* in cui la matrice Hessiana sia non singolare; si può trattare quindi, in particolare, sia di un minimo che di un massimo locale.

## 10.11 Il metodo di Newton con ricerca di linea

Nell'applicazione del metodo di Newton alla minimizzazione non vincolata occorre tener conto dei problemi seguenti:

- (i) la direzione di Newton può non essere definita in  $x^k$  ( $\nabla^2 f(x^k)$  è singolare);
- (ii) la successione prodotta dal metodo di Newton può non essere convergente;
- (iii) si può avere convergenza verso massimi locali.

Per superare tali difficoltà si rende necessario modificare, con opportuni criteri, la direzione di ricerca e introdurre eventualmente anche un parametro scalare da definire attraverso tecniche di ricerca unidimensionale. Le modifiche devono tuttavia essere tali da non far perdere le caratteristiche di rapidità di convergenza proprie del metodo di Newton.

Uno dei criteri più semplici per realizzare una modifica globalmente convergente del metodo di Newton può consistere, ad esempio, nell'utilizzare la direzione dell'antigradiente (o, più in generale, un'opportuna direzione di discesa) quando la direzione di Newton non soddisfa una condizione di sufficiente discesa e nello scegliere  $\alpha^k$  con una tecnica di ricerca unidimensionale tipo-Armijo.

Ad esempio di modifica globalmente convergente del metodo di Newton consideriamo l'algoritmo seguente.

### Metodo di Newton modificato

**Dati.**  $\gamma \in (0, 1/2), \delta \in (0, 1),$ .

**Passo 1.** Si sceglie  $x_0 \in R^n$  e si pone  $k = 0$ .

**Passo 2.** Si calcola  $\nabla f(x^k)$ ; se  $\nabla f(x^k) = 0$  stop.

**Passo 3.** Si calcola  $\nabla^2 f(x^k)$ . Se esiste una soluzione  $s_N$  del sistema lineare:

$$\nabla^2 f(x^k)s = -\nabla f(x^k)$$

e la direzione soddisfa un'opportuna condizione di sufficiente discesa, si pone  $d^k = s_N$ , altrimenti si pone  $d^k = -\nabla f(x^k)$ .

**Passo 4.** Si sceglie  $\alpha^k$  con ricerca di linea di Armijo (se la direzione  $d^k = s_N$ , si può porre  $\Delta^k = 1$ ).

**Passo 5.** Si pone  $x^{k+1} = x^k + \alpha^k d^k$ ,  $k = k + 1$  e si va al passo 2.

**Esempio 21** Sia dato il problema

$$\min -x_1^4 + 2x_1^2 + x_2^2$$

$$\text{Sia } x^0 = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$$

1. Si scriva la prima iterazione del metodo di Newton puro;

2. Dire se il passo  $\alpha = 1$  soddisfa una line search di tipo Armijo utilizzando la direzione di Newton calcolata al passo precedente.

**Soluzione.**

(i) Scriviamo gradiente ed hessiano della funzione obiettivo

$$\nabla f(x) = \begin{pmatrix} -4x_1^3 + 4x_1 \\ 2x_2 \end{pmatrix} \quad \nabla^2 f(x) = \begin{pmatrix} -12x_1^2 + 4 & 0 \\ 0 & 2 \end{pmatrix}$$

La prima iterazione del metodo di Newton puro è:

$$x^1 = x^0 - [\nabla^2 f(x^0)]^{-1} \nabla f(x^0).$$

Calcoliamo i valori di  $\nabla f$  e  $\nabla^2 f$  nel punto  $x^0$

$$\nabla f(x) = \begin{pmatrix} 3/2 \\ 1 \end{pmatrix} \quad \nabla^2 f(x) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

Otteniamo quindi:

$$x^1 = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 3/2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} 3/2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

(ii) Osserviamo preliminarmente che la direzione di Newton ottenuta al passo precedente

$$s_N^0 = - \begin{pmatrix} 3/2 \\ 1/2 \end{pmatrix}$$

è di discesa nel punto  $x^0$ . Infatti risulta  $\nabla f(x^0)^T s_N^0 = -11/4 < 0$  (alla stessa conclusione si poteva arrivare osservando che  $\nabla^2 f(x^0) \succ 0$ ). Posto

$$x^1(\alpha) = x^0 + \alpha s_N^0 = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} - \alpha \begin{pmatrix} 3/2 \\ 1/2 \end{pmatrix} = \begin{pmatrix} 1/2 - \alpha 3/2 \\ 1/2 - \alpha 1/2 \end{pmatrix}$$

dobbiamo verificare se è soddisfatta la condizione

$$f(x^1) \leq f(x^0) + \gamma \alpha \nabla f(x^0)^T s_N^0$$

con  $\alpha = 1$ . Si ottiene

$$x^1(\alpha) = \begin{pmatrix} 1/2 - 3/2 \\ 1/2 - 1/2 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

e quindi  $f(x^1) = 1$ ,  $f(x^0) = 11/16$ . Quindi dobbiamo verificare se risulta

$$1 \leq \frac{11}{16} - \gamma \frac{11}{4}$$

che non è mai soddisfatta qualunque sia il valore di  $\gamma > 0$ .

□

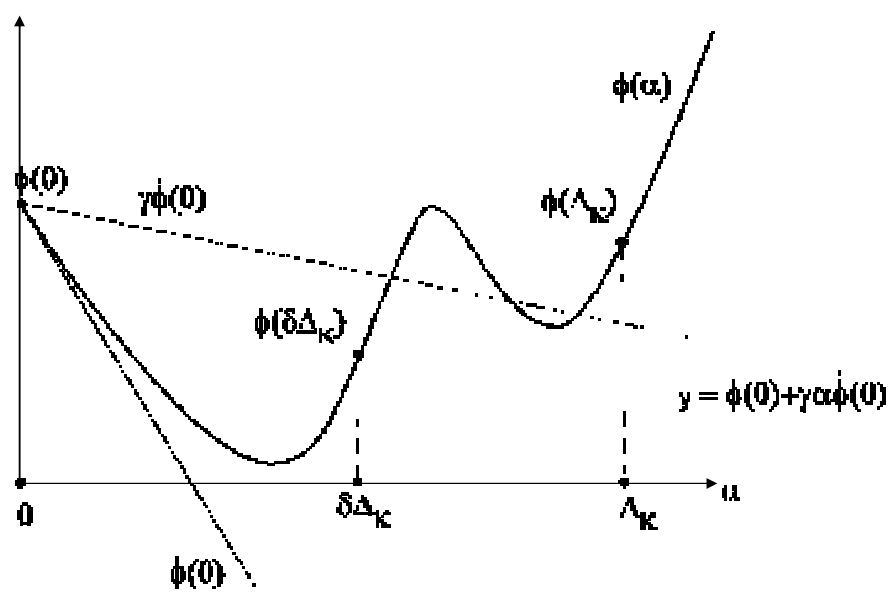


Figure 14: Visione geometrica della condizione di Armijo.

## References

- [1] L.E. Blume and Carl P. Simon “Matematica 2 per l’economia e le Scienze sociali”, a cura di A. Zaffaroni, Università Bocconi Editore, EGEA s.p.a, Ed. italiana 2002.
- [2] D. Bertsekas. “Nonlinear Programming”, Athena Scientific, Belmont, Massachusetts, 1995.
- [3] L. Grippo. “Metodi di ottimizzazione non vincolata”, Rap.Tec. IASI, N. 64, 1988.

## A Richiami sulle norme

Una funzione  $\|\cdot\| : R^n \rightarrow R$  si definisce *norma vettoriale*, se soddisfa le seguenti proprietà:

- $\|v\| \geq 0$  per ogni  $v \in R^n$ ;
- $\|v\| = 0$  se e solo se  $v = 0$ ;
- $\|\alpha v\| = |\alpha| \|v\|$  per ogni  $v \in R^n$  e  $\alpha \in R$ ;
- $\|v + w\| \leq \|v\| + \|w\|$  per ogni  $v, w \in R^n$  (disuguaglianza triangolare).

Le norme più frequentemente utilizzate sono le norme  $\ell_p$  (o norme di Hölder) indicate con  $\|\cdot\|_p$  e definite come

$$\|v\|_p = \left( \sum_{i=1}^n |v_i|^p \right)^{\frac{1}{p}}$$

con  $1 \leq p < \infty$ . Per  $p = 2$  si ottiene la *norma euclidea*

$$\|v\|_2 = \left( \sum_{i=1}^n |v_i|^2 \right)^{\frac{1}{2}}$$

che è quella maggiormente usata. Si può anche definire una norma  $\ell_\infty$  come

$$\|v\|_\infty = \max_{1 \leq i \leq n} |v_i|.$$

Vale sempre la seguente relazione:

$$\|v\|_2 \leq \|v\|_\infty.$$

Nel caso di norma euclidea vale la *disuguaglianza di Schwartz*, ovvero

$$|v^T w| \leq \|v\|_2 \|w\|_2.$$

Consideriamo ora il caso di norme di matrici. Indichiamo con  $M_n$  l'insieme di tutte le matrici  $n \times n$ <sup>3</sup>.

Assegnata una matrice

$$A = (a_{ij})$$

in  $M_n$ , una norma di  $A$  può essere introdotta sia considerando  $A$  come un insieme di  $n^2$  elementi  $a_{ij}$ , sia considerando  $A$  come un operatore lineare da  $R^n$  in  $R^n$ .

Nel primo caso possiamo ovviamente definire come norma di  $A$  una qualsiasi norma vettoriale relativa agli elementi di  $A$ . Un esempio di norma interpretabile come norma del vettore degli elementi è costituito dalla *norma di Frobenius* di  $A$ , data da:

$$\|A\|_F = \left( \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2 \right)^{1/2}$$

---

<sup>3</sup>Ci riferiamo, per semplicità al caso di matrici quadrate, anche se le definizioni espone nel seguito si estendono in modo immediato al caso di matrici rettangolari.

Se  $A, B \in M_n$  si ha:

$$\|A B\|_F \leq \|A\|_F \|B\|_F.$$

Indicando con  $\text{Tr}(A)$  la *traccia* della matrice  $A$ , ossia:

$$\text{Tr}(A) = \sum_{i=1}^n a_{ii},$$

la norma di Frobenius si può anche esprimere nella forma:

$$\|A\|_F = \left( \text{Tr}(A^T A) \right)^{1/2}.$$

Se  $A$  è pensata come un operatore lineare, si può definire la norma di  $A$  ponendo

$$\|A\| = \sup_{x \in R^n, x \neq 0} \frac{\|Ax\|}{\|x\|}, \quad (115)$$

o, equivalentemente:

$$\|A\| = \sup_{\|x\|=1} \|Ax\|.$$

Supponendo, per semplicità, che venga usata la stessa norma vettoriale  $\|\cdot\|$  sia per  $x$  che per  $Ax$ , la norma matriciale definita dalla (115) si dice *norma indotta* dalla norma vettoriale considerata.

Ponendo:

$$\|A\|_p = \sup_{x \in R^n, x \neq 0} \frac{\|Ax\|_p}{\|x\|_p},$$

si ha, in particolare:

$$\|A\|_1 = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|,$$

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|,$$

$$\|A\|_2 = \left( \lambda_M(A^T A) \right)^{1/2},$$

essendo  $\lambda_M(A^T A)$  il massimo autovalore di  $A^T A$ . Se  $A$  è una matrice simmetrica risulta ovviamente

$$\|A\|_2 = \max_{1 \leq i \leq n} |\lambda_i(A)|,$$

essendo  $\lambda_i(A)$  gli autovalori di  $A$ .

Se  $A$  è simmetrica semidefinita positiva, per cui tutti gli autovalori sono non negativi si può porre:

$$\|A\|_2 = \lambda_M(A).$$

## B Richiami sulla differenziazione in $R^n$

Richiamiamo alcuni concetti essenziali sulla differenziazione delle funzioni definite su  $R^n$ , la cui conoscenza è richiesta nello studio dei problemi di ottimizzazione di tipo differenziabile. Nel seguito supponiamo, per semplicità, che  $f$  sia una funzione definita su tutto  $R^n$ ; è immediato tuttavia estendere le definizioni qui riportate al caso in cui  $x$  appartiene ad un insieme aperto  $\mathcal{D} \subseteq R^n$ .

### B.1 Derivate del primo ordine di una funzione reale

Un qualsiasi vettore assegnato  $d \in R^n$  non nullo definisce una *direzione* in  $R^n$ . Una prima nozione di derivata che si può introdurre è quella di *derivata direzionale*.

#### Definizione 9 (Derivata direzionale)

Sia  $f : R^n \rightarrow R$ . Si dice che  $f$  ammette derivata direzionale  $Df(x, d)$  nel punto  $x \in R^n$  lungo la direzione  $d \in R^n$  se esiste finito il limite

$$\lim_{t \rightarrow 0^+} \frac{f(x + td) - f(x)}{t} := Df(x, d). \quad \square$$

Se consideriamo  $f$  come funzione di una sola variabile  $x_j$ , supponendo fissate tutte le altre componenti possiamo introdurre il concetto di *derivata parziale* rispetto a  $x_j$ .

#### Definizione 10 (Derivata parziale)

Sia  $f : R^n \rightarrow R$ . Si dice che  $f$  ammette derivata parziale  $\partial f(x)/\partial x_j$  nel punto  $x \in R^n$  rispetto alla variabile  $x_j$  se esiste finito il limite

$$\lim_{t \rightarrow 0} \frac{f(x_1, \dots, x_j + t, \dots, x_n) - f(x_1, \dots, x_j, \dots, x_n)}{t} := \frac{\partial f(x)}{\partial x_j}. \quad \square$$

L'esistenza della derivata parziale  $\partial f(x)/\partial x_j$  nel punto  $x$  implica, ovviamente, che esistano e coincidano le derivate direzionali lungo le direzioni  $e_j$  e  $-e_j$ .

Se  $f$  ammette derivate parziali rispetto a tutte le componenti possiamo definire il *gradiente* di  $f$  nel punto  $x$ .

#### Definizione 11 (Gradiente)

Sia  $f : R^n \rightarrow R$  ed  $x \in R^n$ . Se esistono le derivate parziali prime di  $f$  in  $x$  definiamo *gradiente* di  $f$  in  $x$  il vettore  $\nabla f(x) \in R^n$

$$\nabla f(x) := \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \dots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix}.$$

A differenza di quanto avviene sulla retta reale, nel caso di  $R^n$  l'esistenza del gradiente non consente, in generale, di poter approssimare, con la precisione voluta, il valore di  $f$  nell'intorno di  $x$  con una funzione lineare dell'incremento. Una tale possibilità è legata ad una nozione più forte di differenziabilità, che è quella riportata nella definizione successiva.

**Definizione 12 (Funzione differenziabile)**

Sia  $f : R^n \rightarrow R$ . Si dice che  $f$  è differenziabile (secondo Frèchet, o in senso forte) nel punto  $x \in R^n$  se esiste  $g(x) \in R^n$  tale che, per ogni  $d \in R^n$  si abbia

$$\lim_{\|d\| \rightarrow 0} \frac{|f(x+d) - f(x) - g(x)^T d|}{\|d\|} = 0,$$

o, equivalentemente, se per ogni  $d \in R^n$  si può porre

$$f(x+d) = f(x) + g(x)^T d + \alpha(x, d),$$

dove  $\alpha(x, d)$  soddisfa:

$$\lim_{\|d\| \rightarrow 0} \frac{\alpha(x, d)}{\|d\|} = 0.$$

L'operatore  $g(x) : R^n \rightarrow R$  si dice derivata (di Frèchet) di  $f$  in  $x$ .  $\square$

È da notare che la sola esistenza di  $\nabla f$  non implica, in generale, la differenziabilità nel senso prima definito. Si dimostra, tuttavia, che se  $\nabla f(x)$  esiste ed è continuo rispetto ad  $x$ , allora  $f$  è differenziabile in  $x$ . (La continuità del gradiente è una condizione *sufficiente* per la differenziabilità in senso forte.) Vale il risultato seguente

**Teorema 48** Sia  $f : R^n \rightarrow R$  e sia  $x \in R^n$ . Si ha:

- (i) se  $f$  è differenziabile in  $x$ , allora  $f$  è continua in  $x$ , esiste il gradiente  $\nabla f(x)$  e  $\nabla f(x)^T$  coincide con la derivata di Frèchet di  $f$  in  $x$
- (ii) se esiste il gradiente  $\nabla f(x)$  e  $\nabla f(x)$  è continuo rispetto a  $x$ , allora  $f$  è differenziabile in  $x$ , e la derivata di Frèchet di  $f$  in  $x$  coincide con  $\nabla f(x)^T$ .  $\square$

Dal teorema precedente segue che se  $\nabla f(x)$  è continuo si può scrivere, per ogni  $d \in R^n$ :

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

dove  $\alpha(x, d)$  soddisfa:

$$\lim_{\|d\| \rightarrow 0} \frac{\alpha(x, d)}{\|d\|} = 0.$$

Se  $f$  è differenziabile, è immediato verificare che esiste anche la derivata direzionale di  $f$  lungo una qualsiasi direzione  $d \in R^n$  e risulta:

$$\lim_{t \rightarrow 0^+} \frac{f(x+td) - f(x)}{t} = \nabla f(x)^T d.$$

## B.2 Differenziazione di un vettore di funzioni

Sia  $g : R^n \rightarrow R^m$  un vettore a  $m$  componenti di funzioni reali. Possiamo introdurre la definizione seguente, che estende la nozione di gradiente.

### Definizione 13 (Matrice Jacobiana)

Sia  $g : R^n \rightarrow R^m$  e  $x \in R^n$ . Se esistono le derivate parziali prime  $\partial g_i(x)/\partial x_j$ , per  $i = 1 \dots, m$  e  $j = 1 \dots, n$  in  $x$  definiamo matrice Jacobiana di  $g$  in  $x$  la matrice  $m \times n$

$$J(x) := \begin{pmatrix} \frac{\partial g_1(x)}{\partial x_1} & \cdots & \frac{\partial g_1(x)}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial g_m(x)}{\partial x_1} & \cdots & \frac{\partial g_m(x)}{\partial x_n} \end{pmatrix}.$$

Possiamo estendere in modo ovvio la nozione di differenziabilità al caso di un vettore di funzioni.

### Definizione 14 (Derivata prima di un vettore di funzioni)

Sia  $g : R^n \rightarrow R^m$  e  $x \in R^n$ . Si dice che  $g$  è differenziabile (secondo Frèchet, o in senso forte) nel punto  $x \in R^n$  se esiste una matrice  $G(x)$  tale che, per ogni  $d \in R^n$  si abbia

$$\lim_{\|d\| \rightarrow 0} \frac{\|g(x+d) - g(x) - G(x)d\|}{\|d\|} = 0,$$

o, equivalentemente, se per ogni  $d \in R^n$  si può porre

$$g(x+d) = g(x) + G(x)d + \gamma(x, d),$$

dove  $\gamma(x, d)$  soddisfa:

$$\lim_{\|d\| \rightarrow 0} \frac{\gamma(x, d)}{\|d\|} = 0.$$

L'operatore lineare  $G(x) : R^n \rightarrow R^m$  si dice derivata (di Frèchet) di  $g$  in  $x$ .  $\square$

Anche in questo caso, ovviamente, la sola esistenza della matrice Jacobiana in  $x$  non implica la differenziabilità e si ha il teorema seguente.

**Teorema 49** Sia  $g : R^n \rightarrow R^m$  e  $x \in R^n$ . Si ha:

- (i) se  $g$  è differenziabile in  $x$ , allora  $g$  è continua in  $x$ , esiste la matrice Jacobiana  $J(x)$  e  $J(x)$  coincide con la derivata di Frèchet di  $g$  in  $x$
- (ii) se esiste la matrice Jacobiana  $J(x)$  di  $g$  in  $x$  e  $J(x)$  è continua rispetto a  $x$ , allora  $g$  è differenziabile in  $x$ , e la derivata di Frèchet di  $g$  in  $x$  coincide con  $J(x)$ .  $\square$

Dal teorema precedente segue che se  $J(x)$  è continua si può scrivere, per ogni  $d \in R^n$ :

$$g(x+d) = g(x) + J(x)d + \gamma(x, d),$$

dove  $\gamma(x, d)$  soddisfa:

$$\lim_{\|d\| \rightarrow 0} \frac{\gamma(x, d)}{\|d\|} = 0.$$

Se  $m = 1$  si ha ovviamente  $J(x) = \nabla g(x)^T$ .

Per analogia con la notazione usata per il gradiente si può usare anche la notazione  $\nabla g(x)^T$  per indicare la derivata prima di  $g$ , ossia

$$\nabla g(x) = J(x)^T = (\nabla g_1(x), \dots, \nabla g_m(x)).$$

### B.3 Derivate del secondo ordine di una funzione reale

Sia  $f : R^n \rightarrow R$  una funzione reale. Con riferimento alle derivate del secondo ordine, possiamo introdurre innanzitutto la definizione seguente.

#### Definizione 15 (Matrice Hessiana)

Sia  $f : R^n \rightarrow R$  e  $x \in R^n$ . Se esistono le derivate parziali seconde  $\partial^2 f(x)/\partial x_i \partial x_j$ , per  $i = 1 \dots, n$  e  $j = 1 \dots, n$  in  $x$  definiamo matrice Hessiana di  $f$  in  $x$  la matrice  $n \times n$

$$\nabla^2 f(x) := \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix}.$$

Possiamo introdurre la nozione di derivata seconda mediante la definizione seguente.

#### Definizione 16 (Derivata seconda)

Sia  $f : R^n \rightarrow R$  e  $x \in R^n$ . Si dice che  $f$  è due volte differenziabile (secondo Frèchet, o in senso forte) nel punto  $x \in R^n$  se la derivata prima di  $f$ ,  $\nabla f(x)^T$  è differenziabile in  $x$ . La derivata prima di  $\nabla f(x)^T$  si dice derivata seconda (di Frèchet) di  $f$ .  $\square$

Anche in questo caso, la sola esistenza della matrice Hessiana (che possiamo interpretare come matrice Jacobiana di  $\nabla f(x)$ ) non implica la differenziabilità e si ha il teorema seguente.

**Teorema 50** Sia  $f : R^n \rightarrow R$  e  $x \in R^n$ . Si ha:

- (i) se  $f$  è due volte differenziabile in  $x$ , allora il gradiente  $\nabla f(x)$  esiste ed è continuo in  $x$ , la matrice Hessiana  $\nabla^2 f(x)$  esiste ed è una matrice simmetrica e  $\nabla^2 f(x)$  coincide con la derivata seconda di Frèchet di  $f$  in  $x$
- (ii) se esiste la matrice Hessiana  $\nabla^2 f(x)$  in  $x$  e  $\nabla^2 f(x)$  è continua rispetto a  $x$ , allora  $f$  è due volte differenziabile in  $x$ ,  $\nabla^2 f(x)$  è necessariamente simmetrica e la derivata seconda di Frèchet di  $f$  in  $x$  coincide con  $\nabla^2 f(x)$ .  $\square$

Dal teorema precedente segue che se  $\nabla^2 f(x)$  è continua si può scrivere, per ogni  $d \in R^n$ :

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

dove  $\beta(x, d)$  soddisfa:

$$\lim_{\|d\| \rightarrow 0} \frac{\beta(x, d)}{\|d\|^2} = 0.$$

Nelle stesse ipotesi si ha anche, come si è detto, che la matrice Hessiana  $\nabla^2 f(x)$  è una **matrice simmetrica**, ossia si ha, per  $i, j = 1 \dots, n$ :

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{\partial^2 f(x)}{\partial x_j \partial x_i}.$$

Osserviamo anche che se  $\nabla^2 f(x)$  è continua si può scrivere

$$\nabla f(x+d) = \nabla f(x) + \nabla^2 f(x)d + \gamma(x,d),$$

con

$$\lim_{\|d\| \rightarrow 0} \frac{\gamma(x,d)}{\|d\|} = 0.$$

#### B.4 Teorema della media e formula di Taylor

Nel caso di funzioni differenziabili valgono anche i risultati seguenti (che si possono tuttavia stabilire sotto ipotesi più deboli).

##### **Teorema 51 (Teorema della Media)**

*Sia  $f : R^n \rightarrow R$  una funzione differenziabile. Allora, per ogni  $h \in R^n$ , si può scrivere*

$$f(x+h) = f(x) + \nabla f(z)^T h,$$

*in cui  $z \in R^n$  è un punto opportuno (dipendente da  $x$  e  $h$ ) tale che  $z = x + \zeta h$ , con  $\zeta \in (0,1)$ . □*

Possiamo dare di tale risultato anche una formulazione integrale.

**Teorema 52** *Sia  $f : R^n \rightarrow R$  una funzione differenziabile. Allora, per ogni  $h \in R^n$ , si può scrivere*

$$f(x+h) = f(x) + \int_0^1 \nabla f(x+th)^T h dt, \quad \square$$

Utilizzando le derivate seconde si ha il risultato seguente.

##### **Teorema 53 Teorema di Taylor**

*Sia  $f : R^n \rightarrow R$  una funzione due volte differenziabile. Allora, per ogni  $h \in R^n$  si può scrivere:*

$$f(x+h) = f(x) + h^T \nabla f(x) + \frac{1}{2} h^T \nabla^2 f(w) h$$

*in cui  $w \in R^n$  è un punto opportuno (dipendente da  $x$  e  $h$ ) tale che  $w = x + \xi h$ , con  $\xi \in (0,1)$ . □*

Anche in questo caso possiamo considerare una formulazione di tipo integrale.

**Teorema 54** *Sia  $f : R^n \rightarrow R$  una funzione due volte differenziabile. Allora, per ogni  $h \in R^n$  si può scrivere:*

$$f(x+h) = f(x) + h^T \nabla f(x) + \int_0^1 (1-t) h^T \nabla^2 f(x+th) h dt.$$

□

Nel caso di funzioni vettoriali  $g : R^n \rightarrow R^m$  non è possibile stabilire un teorema della media. È tuttavia possibile considerare un'espressione di tipo integrale.

**Teorema 55** Sia  $g : R^n \rightarrow R^m$  una funzione differenziabile. Allora, per ogni  $h \in R^n$ , si può scrivere

$$g(x+h) = g(x) + \int_0^1 J(x+th)h dt,$$

in cui  $J$  è la matrice Jacobiana di  $g$ .  $\square$

Come caso particolare del risultato precedente, se  $g$  è il gradiente  $\nabla f$  di una funzione due volte differenziabile  $f : R^n \rightarrow R$ , si ha:

$$\nabla f(x+h) = \nabla f(x) + \int_0^1 \nabla^2 f(x+th)h dt.$$

### Esempi

Alcuni esempi di interesse di funzioni differenziabili sono:

$$f(x) = c^T x; \text{ si ha: } \nabla f(x) = c, \quad \nabla^2 f(x) = 0;$$

$$f(x) = \frac{1}{2}x^T Qx + c^T x, \text{ (con } Q \text{ simmetrica); si ha: } \nabla f(x) = Qx + c; \quad \nabla^2 f(x) = Q;$$

$$f(x) = \|Ax - b\|^2; \text{ si ha: } \nabla f(x) = 2A^T(Ax - b); \quad \nabla^2 f(x) = 2A^T A$$