Appunti delle esercitazioni di Ricerca Operativa

a cura di P. Detti * e G. Ciaschetti *

1 Esercizi sulle condizioni di ottimalità per problemi di ottimizzazione non vincolata

Esempio 1 *Sia data la funzione*

$$f(x_1, x_2, x_3) = 4x_1^2 + 2x_2^2 + x_3^2 - x_1x_2 + x_2x_3 - 5x_1 - 9x_2 + x_3$$

studiare l'esistenza di punti di minimo.

La funzione è continua con derivate parziali del primo e del secondo ordine continue. Applichiamo le condizioni necessarie del primo ordine per determinare l'insieme dei punti stazionari $PS = \{x : x \in R^n, \nabla f(x) = 0\}$. Ponendo $\nabla f(x) = 0$, si ottiene il seguente sistema di equazioni:

$$\begin{aligned} 8x_1 - x_2 - 5 &= 0 \\ -x_1 + 4x_2 + x_3 - 9 &= 0 \\ x_2 + 2x_3 + 1 &= 0 \end{aligned}$$

da cui si ottiene l'unica soluzione $x_1 = 1$, $x_2 = 3$, $x_3 = -2$. Consideriamo le condizioni necessarie del secondo ordine. La *matrice Hessiana* è in questo caso costante ed è pari a:

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{bmatrix} = \begin{bmatrix} 8 & -1 & 0 \\ -1 & 4 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

$\nabla^2 f(x)$ è inoltre definita positiva, ha, infatti, i determinanti di tutti i minori principali positivi:

$$8 > 0, \quad \begin{vmatrix} 8 & -1 \\ -1 & 4 \end{vmatrix} = 31 > 0, \quad |\nabla^2 f(x)| = 54 > 0$$

*Dipartimento di Ingegneria dell'Informazione - Università di Siena

La funzione in esame è quindi convessa ovunque, ed il punto $x_1 = 1$, $x_2 = 3$, $x_3 = -2$ è un punto di minimo globale stretto.

Esempio 2 *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.

La funzione è coerciva. Infatti, possiamo scrivere

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

e poichè

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

si ha

$$\lim_{||x|| \to \infty} f(x_1, x_2) = +\infty$$

Poichè la funzione è continua, esiste almeno un punto stazionario che è un punto di minimo globale.

La funzione è continua con derivate parziali del primo e del secondo ordine continue. Applichiamo le condizioni necessarie del primo ordine ($\nabla f(x) = 0$) per determinare l'insieme dei punti stazionari. Si ha il sistema:

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

da cui si ottengono le soluzioni: $A = (0, 0)$, $B = (\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2})$, $C = (-\frac{\sqrt{3}}{2}, -\frac{\sqrt{3}}{2})$.

Per determinare la natura dei punti A, B e C imponiamo le condizioni del secondo ordine. La matrice Hessiana di f è:

$$\nabla^2 f(x) = \begin{bmatrix} 12x_1^2 & -3 \\ -3 & 12x_2^2 \end{bmatrix}$$

Notare che $\nabla^2 f(x)$ non è definita positiva ovunque (basta scegliere $x_1 = x_2 = 0$) e quindi la funzione considerata non è convessa.

Calcoliamo la matrice Hessiana nei punti stazionari.

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

In A l'Hessiana è indefinita. Infatti, né $\nabla^2 f(0, 0)$ né $-\nabla^2 f(0, 0)$ sono matrici definite o semidefinite postive (e, quindi, $\nabla^2 f(0, 0)$ non è né definita o semidefinita positiva, né definita o semidefinita negativa). Si tratta quindi di un punto di *sella* con valore della funzione $f(0,0) = 0$. Nei punti B e C si ha:

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

che è una matrice definita positiva con autovalori $\lambda_{min} = 2$, $\lambda_{max} = 4$. B e C sono quindi due punti di minimo locale stretto con valore della funzione obiettivo $f(B) = f(C) = -\frac{9}{8}$.

Poichè la funzione in esame è radialmente illimitata (coerciva) e non esistono altri punti stazionari con valore inferiore della funzione obiettivo, B e C sono anche minimi globali.

Esempio 3 *Sia data la funzione*

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

studiare l'esistenza di punti di minimo.

La funzione non è coerciva. Infatti, lungo la direzione definita da $x_2 = \bar{x}_2 = cost$ si ha

$$\lim_{\substack{x_1 \to -\infty \\ ||x|| \to \infty}} f(x_1, x_2) = -\infty$$

In particolare, poichè esiste una direzione lungo cui la funzione è illimitata inferiormente, non esiste un punto di minimo globale.

Applichiamo le condizioni del primo ordine. Il gradiente della funzione è:

$$\nabla f = 3 \begin{bmatrix} x_1^2 - 4x_2 \\ -4x_1 + 8x_2^2 \end{bmatrix}$$

I punti stazionari sono individuati risolvendo il sistema:

$$\begin{cases} x_1^2 - 4x_2 = 0 \\ -x_1 + 2x_2^2 = 0 \end{cases}$$

da cui si ottengono i punti $A = (0,0)$ e $B = (2,1)$. Calcoliamo la matrice Hessiana nei punti A e B.

$$\nabla^2 f(0,0) = 6 \begin{bmatrix} 0 & -2 \\ -2 & 0 \end{bmatrix}$$

In A l'Hessiana è indefinita con autovalori $\lambda_{min} = -2$, $\lambda_{max} = 2$. Si tratta quindi di un punto di sella.

$$\nabla^2 f(2,1) = 12 \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix}$$

In B l'Hessiana è definita positiva (i minori principali sono tutti positivi) con autovalori $\lambda_{min} = 0,69$, $\lambda_{max} = 4,3$. Si tratta quindi di un punto di minimo locale stretto.

Esempio 4 *Siano date le funzioni*

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

studiare la natura del punto $(0, 0)$.

Calcoliamo i gradienti e le matrici Hessiane di f e g.

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

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

Il punto $(0, 0)$ soddisfa le condizioni necessarie del secondo ordine, ma poichè le matrici Hessiane sono semidefinite positive le condizioni non sono sufficienti. Infatti risulta

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

Per quanto riguarda la funzione $g(x)$, si noti che $x^T \nabla^2 g(x) x \geq 0$ per ogni $x \in R^2$. La matrice Hessiana di $g(x)$ è quindi semidefinita positiva ovunque. Si può dimostrare che una funzione due volte continuamente differenziabile è convessa se e solo se la sua matrice hessiana è sempre semidefinita positiva. Possiamo concludere, quindi, che la $g(x)$ è convessa e che il punto $(0, 0)$ è un punto di minimo globale (per funzioni convesse le condizioni di ottimalità del primo ordine sono condizioni sufficienti di ottimalità globale). Consideriamo la funzione $f(x)$. Le considerazioni sopra riportate non possono essere applicate in questo caso. Per determinare la vera natura del punto $(0, 0)$ consideriamo lo sviluppo di Taylor della funzione in un punto $\beta \in R^2$ nell'intorno del punto $(0, 0)$. In generale si ha:

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

dove $\bar{x}$ è ora un punto del segmento che unisce x a $x + h$.

Poiché $f(0, 0) = 0$ e $\nabla f(0, 0) = (0 \quad 0)^T$, si ottiene il seguente sviluppo

$$f(0 + \beta) = \frac{1}{2} \beta^T \nabla^2 f(y) \beta$$

dove $y = \theta\beta$ con $\theta \in (0, 1)$ e β è un vettore con componenti (β_1, β_2). Si ha che tutti i punti $\beta = (\beta_1, \beta_2)$ con compenente β_2 negativa, rendono l'Hessiana indefinita. Il punto $(0, 0)$ è quindi un punto di sella per f.

Esempio 5 *Consideriamo la funzione di Rosembrock*

$$f(x_1, x_2) = c(x_2 - x_1^2)^2 + (1 - x_1)^2$$

con $c > 0$. *studiare l'esistenza di punti di minimo.*

La funzione è coerciva. Infatti risulta $f(x_1, x_2) \geq 0$ e inoltre non esiste una direzione lungo la quale la f si mantenga costante.

Calcoliamo il gradiente e la matrice Hessiana.

$$\nabla f = \begin{bmatrix} -4cx_1(x_2 - x_1^2) - 2(1 - x_1) \\ 2c(x_2 - x_1^2) \end{bmatrix}$$

$$\nabla^2 f = 2 \begin{bmatrix} 2c(3x_1^2 - x_2) + 1 & -2cx_1 \\ -2cx_1 & c \end{bmatrix}$$

I punti stazionari ($\nabla f = 0$) si ottengono risolvendo il seguente sistema:

$$\begin{cases} -4cx_1(x_2 - x_1^2) - 2(1 - x_1) = 0 \\ 2c(x_2 - x_1^2) = 0 \end{cases} \qquad \begin{cases} 1 - x_1 = 0 \\ x_2 = x_1^2 \end{cases} \begin{cases} x_1 = 1 \\ x_2 = 1 \end{cases}$$

Nel punto $(1, 1)$ si ha:

$$\nabla^2 f(1, 1) = \begin{bmatrix} 4c + 1 & -2c \\ -2c & c \end{bmatrix}$$

La matrice Hessiana è definita positiva per $c > 0$. Il punto $(1, 1)$ è quindi un punto di minimo globale.

Esempio 6 *Dato un insieme di punti di coordinate* (x_i, y_i), $i = 1, \dots, n$, *trovare la retta* $y = ax + b$ *che* meglio approssima *la relazione tra* x *ed* y.

Uno dei metodi utilizzati per determinare la retta richiesta è quello conosciuto con il nome di *approssimazione ai minimi quadrati*. Tale metodo consiste nello scegliere a e b in modo da minimizzare la somma dei quadrati delle seguenti quantità

$$e_i = y_i - (ax_i + b) \quad i = 1, \dots, n$$

in cui, per ogni i, $ax_i + b$ e y_i sono detti rispettivamente valore atteso e valore attuale di y_i. La funzione, che deve essere minimizzata, è quindi:

$$f(a, b) = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - ax_i - b)^2$$

Imponiamo le condizioni necessarie del primo ordine ($\nabla f = 0$) per determinare l'insieme dei punti stazionari.

$$\begin{cases} \frac{\partial f}{\partial a} = -2 \sum_{i=1}^{n} (y_i - ax_i - b)x_i = 0 \\ \frac{\partial f}{\partial b} = -2 \sum_{i=1}^{n} (y_i - ax_i - b) = 0 \end{cases} \qquad \begin{cases} (\sum_{i=1}^{n} x_i^2)a + (\sum_{i=1}^{n} x_i)b = \sum_{i=1}^{n} x_i y_i \\ (\sum_{i=1}^{n} x_i)a + nb = \sum_{i=1}^{n} y_i \end{cases}$$

Dividendo la seconda equazione per n, e ponendo $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ e $\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$ si ottiene:

$$a\bar{x} + b = \bar{y}$$

La retta cercata passa quindi per il punto $(\bar{x}, \bar{y})$ che è noto. Moltiplicando la seconda equazione per $\bar{x}$ e sottraendola alla prima equazione si ottiene:

$$(\sum_{i=1}^{n} x_i^2 - \bar{x} \sum_{i=1}^{n} x_i)a = \sum_{i=1}^{n} x_i y_i - \bar{x} \sum_{i=1}^{n} y_i$$

da cui

$$a = \frac{\sum_{i=1}^{n} x_i y_i - \bar{x} \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i^2 - \bar{x} \sum_{i=1}^{n} x_i}$$

e

$$b = \bar{y} - a\bar{x}$$

La matrice Hessiana è:

$$\begin{bmatrix} (\sum_{i=1}^{n} x_i^2) & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & n \end{bmatrix}$$

Tale matrice ha l'elemento $a_{1,1}$ sicuramente positivo perchè somma di quadrati, ed ha determinate pari a

$$n(\sum_{i=1}^{n} x_i^2) - (\sum_{i=1}^{n} x_i)^2$$

che, sviluppando il secondo termine, può essere scritto come

$$n(\sum_{i=1}^{n} x_i^2) - \sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} \sum_{j=i+1}^{n} x_i x_j) = (n-1)(\sum_{i=1}^{n} x_i^2) - \sum_{i=1}^{n} \sum_{j=i+1}^{n} x_i x_j =$$

$$= \sum_{i=1}^{n} \sum_{j=i+1}^{n} (x_i - x_j)^2$$

La matrice Hessiana è quindi definita positiva e costante. Di conseguenza la funzione è convessa ed il punto (a, b) trovato è un punto di minimo globale.

Esempio 7 *Un venditore deve pianificare per il nuovo anno gli ordini di approvvigionamento presso il suo fornitore di un totale di D unità di un determinato bene (da acquistare durante l'anno presso un fornitore). Nel nuovo anno, si assume che tale quantità sarà venduta dal venditore ad un tasso costante. L'obiettivo è determinare un piano di ordini, decidere cioé quando e quanto ordinare durante il nuovo anno, in modo da (I) ordinare complessivamente D e (II) minimizzare il costo totale di immagazinamento e di ordinazione. Si supponga che il tempo che intercorre dall'emissione di un ordine all'arrivo effettivo del bene sia trascurabile, e che un ordine possa essere emesso, quindi, quando il magazzino è vuoto.*

Siano h il costo unitario di immagazzinamento ed A il costo di emissione di un ordine (indipendente dalla quantità ordinata). Si indichi con x la quantità di bene ordinata ogni volta che viene emesso un ordine. Si noti che, poiché h ed A sono costanti nell'anno, non conviene ordinare quantità diverse in ordini diversi. L'intervallo di tempo t che intercorre tra due ordini successivi e la quantità ordinata x sono legati dalla relazione

$$t = \frac{x}{D}$$

Infatti, poichè il tasso di vendita è costante, D può essere visto come il tasso di vendita del bene nell'unità di tempo (in questo caso un anno). Siano rispettivamente t_1 e t_2 l'istante di tempo in cui è avvenuto l'ultimo ordine, e l'istante di tempo immediatamente successivo a t_1 in cui il magazzino è vuoto (Figura 1). La quantità del bene che deve essere ordinata in t_2 per riportare il magazzino al livello avuto in t_1 è $x = D(t_2 - t_1)$, dove $t_2 - t_1 = t$.

Per soddisfare la domanda D durante il nuovo anno, D/x ordini dovranno essere emessi. Il costo totale di ordinazione è quindi:

$$\frac{AD}{x}$$

Poichè il tasso di vendita del bene è costante, il livello medio del magazzino è $x/2$, e quindi il costo totale di immagazzinamento è:

$$\frac{hx}{2}$$

Il costo totale che si vuole minimizzare è quindi dato dalla seguente funzione:

$$f(x) = \frac{hx}{2} + \frac{AD}{x}$$

La quantitá x si ottiene annullando la derivata prima di $f(x)$:

$$f'(x) = \frac{h}{2} - \frac{AD}{x^2} = 0$$

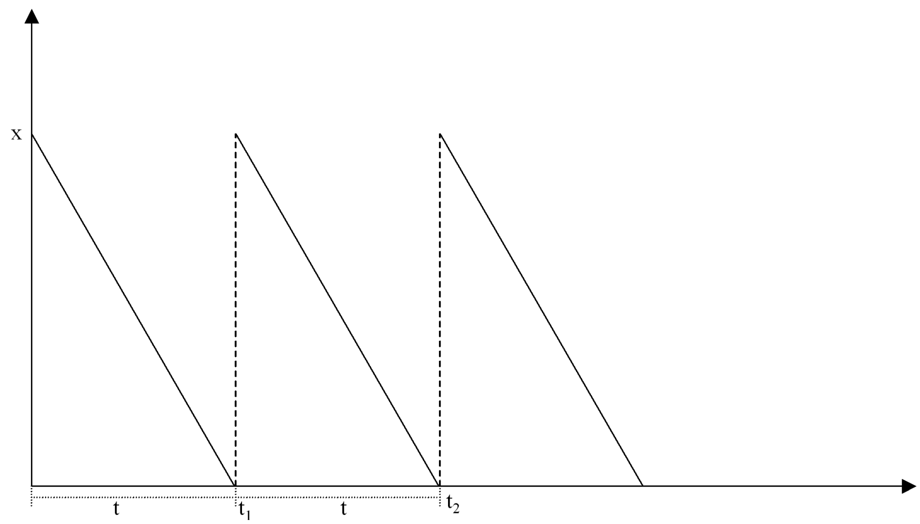

Figure 1: Pianificazione degli ordini.

da cui

$$x^* = \sqrt{AD/h}$$

Notare che la x è una quantità positiva. D'altra parte una x negativa non avrebbe senso per il problema in esame.
La derivata secona è:

$$f''(x) = \frac{2AD}{x^3}$$

che è sempre positiva se $x \geq 0$. La funzione è quindi una funzione convessa per $x \geq 0$ e x^* è un punto di minimo globale.

Esempio 8 *Consideriamo il problema riportato nell'esempio precedente con l'ipotesi aggiuntiva che il bene possa essere venduto anche se il magazzino è fisicamente vuoto. In altre parole il venditore può vendere il bene anche se questo non è in magazzino (in inglese questa situazione è indicata con il termine* shortage*). Si suppone, però, che per ogni unità di merce venduta, ma non fisicamente consegnata al cliente finale, debba essere sostenuto un determinato costo (costo di* shortage*). Determinare un piano di ordini del bene in modo tale che il costo totale di immagazinamento, di ordinazione e di shortage sia minimo.*

Siano h il costo unitario di immagazzinamento per unità di tempo, A il costo di emissione di un ordine (indipendente dalla quantità ordinata), e b il costo unitario di shortage per unità di tempo. Si indichi con x la quantità di bene ordinata ogni volta che viene emesso un ordine. L'intervallo tra due ordini successivi t è

$$t = \frac{x}{D}$$

Sia t_s l'intervallo di t in cui si ha shortage, e sia $t_i = t - t_s$ l'intervallo di tempo in cui il magazzino non è vuoto.

Per soddisfare la domanda D durante il nuovo anno, D/x ordini dovranno essere emessi.

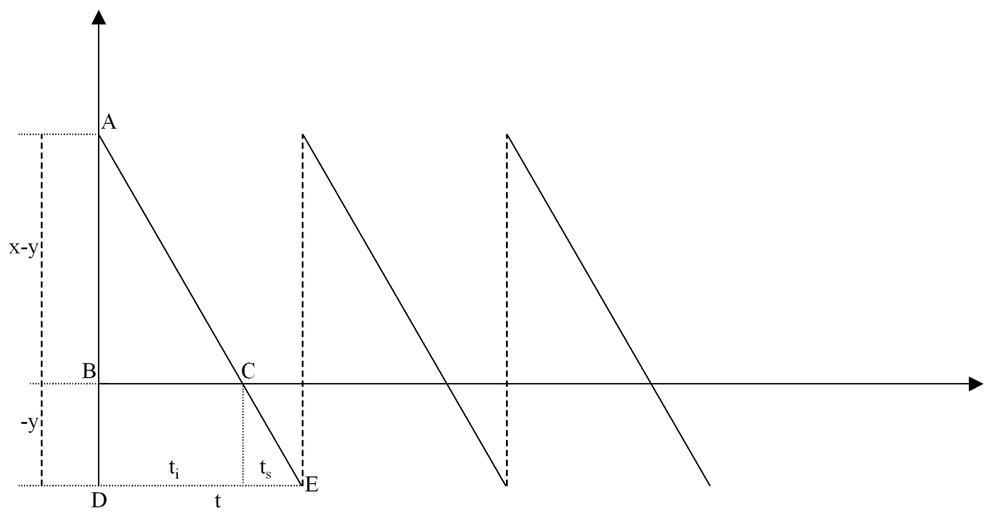

Figure 2: Pianificazione degli ordini con shortage.

Il costo totale di ordinazione è quindi:

$$\frac{AD}{x}$$

Sia $\frac{t_i}{t}$ la frazione di t in cui non si ha shortage. Poichè il tasso di vendita del bene è costante, il livello medio del magazzino in tale frazione è $(x-y)/2$, e il costo di immagazzinamento è quindi:

$$h\frac{(x-y)}{2}\frac{t_i}{t}$$

Dalla similitudine dei triangoli ABC e ADE di Figura 2 discende la relazione:

$$\frac{t_i}{t} = \frac{x-y}{x}$$

e quindi il costo di immagazzinamento diventa:

$$h\frac{(x-y)^2}{2x}$$

Nella frazione di tempo $\frac{t_s}{t}$ si ha uno shortage medio di $y/2$ unità. Il costo di shortage è quindi:

$$b\frac{y}{2}\frac{t_s}{t}$$

Dalla similitudine dei triangoli ADE e CFE di Figura 2 discende la relazione:

$$\frac{t_s}{t} = \frac{y}{x}$$

per cui il costo di shortage è:

$$b\frac{y^2}{2x}$$

Il costo totale che si vuole minimizzare è quindi dato dalla seguente funzione:

$$f(x,y) = \frac{AD}{x} + h\frac{(x-y)^2}{2x} + b\frac{y^2}{2x}$$

Applichiamo le condizioni necessarie del primo ordine per determinarte l'insieme dei punti stazionari.

$$\nabla f = \begin{bmatrix} -\frac{AD}{x^2} + \frac{h}{2} - h\frac{y^2}{2x^2} - b\frac{y^2}{2x^2} \\ -h + \frac{(h+b)y}{x} \end{bmatrix} = 0$$

da cui si ottiene:

$$x^* = \sqrt{2AD}\sqrt{\frac{(h+b)}{hb}}$$

$$y^* = \sqrt{2AD}\sqrt{\frac{h}{(h+b)b}}$$

Consideriamo le condizioni necessarie del secondo ordine. La matrice Hessiana è:

$$\begin{bmatrix} 2\frac{AD}{x^3} + \frac{(h+b)y^2}{x^3} & \frac{-(h+b)y}{x^2} \\ \frac{-(h+b)y}{x^2} & \frac{h+b}{x} \end{bmatrix}$$

Tale matrice per $x, y \geq 0$ è definita positiva e quindi il punto (x^*, y^*) è un minimo globale per la $f(x,y)$ in R^{2^+}.

Esempio 9 *Si determinino i punti di minimo della funzione:* $f(x) = x_1^3 + x_2^3 - 4x_1x_2$. *Dire inoltre se esiste un punto di minimo globale.*

La funzione è continua e differenziabile due volte con continuitá in $\mathcal{R}^2$. Imponendo le condizioni necessarie del 1° ordine, individuiamo l'insieme dei punti stazionari.

$$\nabla f(x) = (3x_1^2 - 4x_2, 3x_2^2 - 4x_1)^T = 0$$

da cui otteniamo il sistema

$$\begin{cases} 3x_1^2 - 4x_2 = 0 \\ 3x_2^2 - 4x_1 = 0 \end{cases}$$

Risolvendo, troviamo i punti $A = (0,0), B = (\frac{4}{3}, \frac{4}{3})$. Per caratterizzare i punti A e B, studiamo ora la matrice hessiana $\nabla^2 f(x)$.

$$\nabla^2 f(x) = [\begin{matrix} 6x_1 & -4 \\ -4 & 6x_2 \end{matrix}]$$

$\nabla^2 f(x)$ non è ovunque definita positiva. Nel punto A, si ha

$$\nabla^2 f(A) = [\begin{array}{cc} 0 & -4 \\ -4 & 0 \end{array}]$$

che non è semidefinita positiva, e dunque, il punto A non è un punto di minimo. Poiché $-\nabla^2 f(A)$ non è semidefinita positiva, si tratta di un punto di sella. Per quanto riguarda il punto B, abbiamo

$$\nabla^2 f(B) = [\begin{array}{cc} 8 & -4 \\ -4 & 8 \end{array}]$$

che è definita positiva. Dunque, B è un punto di minimo locale.
Ci chiediamo ora se esistono punti di minimo globale.
Poichè risulta $\lim_{x_1 \to -\infty} f(x) = \lim_{x_2 \to -\infty} f(x) = -\infty$ possiamo concludere allora che non esiste un punto di minimo globale.

Esempio 10 *Si trovino i punti di minimo della funzione* $f(x) = x_1^2 - 3x_2^2 - 2x_1x_2 + 4x_1$

La funzione è continua e differenziabile due volte con continuitá in $\mathcal{R}^2$. Applicando le condizioni necessarie del primo ordine, cioè imponendo $\nabla f(x) = 0$, abbiamo il sistema

$$\begin{cases} 2x_1 - 2x_2 + 4 = 0 \\ -6x_2 - 2x_1 = 0 \end{cases}$$

da cui otteniamo l'unico punto $A = (-\frac{3}{2}, \frac{1}{2})$. La matrice hessiana è data da

$$\nabla^2 f() = [\begin{array}{cc} 2 & -2 \\ -2 & -6 \end{array}]$$

Essa non è definita positiva né semidefinita posivita. Studiamo $-\nabla^2 f(x)$ per vedere se è semidefinita negativa:

$$-\nabla^2 f() = [\begin{array}{cc} -2 & +2 \\ +2 & +6 \end{array}]$$

non è semidefinita positiva, dunque, si può concludere che l'hessiana è indefinita, e il punto A è un punto di sella.

2 Algoritmi di ottimizzazione non vincolata

Esempio 11 *Sia data la funzione*

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

- *studiare l'esistenza di punti di minimo;*
- *a partire dal punto iniziale*

$$x^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

  - *applicare il metodo del gradiente;*
  - *applicare il metodo di Newton.*

La funzione in esame è una funzione quadratica.
Applichiamo le condizioni del primo ordine.

$$\nabla f = \begin{bmatrix} 4x_1 + 2x_2 + 1 \\ 2x_1 + 2x_2 - 1 \end{bmatrix}$$

Da cui il sistema:

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

Risolvendo si ottiene la soluzione:

$$x^* = \begin{bmatrix} -1 \\ 3/2 \end{bmatrix}$$

con valore $f(x^*) = -5/4$.
Calcoliamo la matrice Hessiana.

$$\nabla^2 f(x) = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}$$

L'Hessiana è costante e definita positiva, infatti gli autovalori sono $\lambda_M = 3 + \sqrt{5}$ e $\lambda_m = 3 - \sqrt{5}$. La funzione è quindi convessa e x^* è un punto di minimo globale.

Metodo del gradiente
Applichiamo l'algoritmo del gradiente a partire dal punto x^0 e consideriamo per la ricerca del passo α sia una ricerca esatta (ossia scegliendo α in modo da minimizzare la funzione $\phi(\alpha) = f(x^{(k)} + \alpha d^{(k)})$), sia il metodo di Armijo. Notare che poichè la funzione è quadratica è possibile calcolare il passo con la ricerca esatta in modo analitico. Infatti risulta:

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

e quindi annullando la derivata rispetto ad α si ottiene $\nabla f(x)^T d + \alpha d^T \nabla f(x)^2 d = 0$. Se $\nabla f(x)^2$ è definita positiva, si ha:

$$\alpha^{(k)} = -\frac{\nabla f(x^{(k)})^T d^{(k)}}{d^{k^T} \nabla f(x^{(k)})^2 d^{(k)}}$$

Nel punto iniziale $x^{(0)}$:

$$\nabla f(x^{(0)}) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \neq 0$$

e quindi non è un punto stazionario.

Iterazione 0

La direzione di discesa è:

$$d^{(0)} = -\nabla f(x^{(0)}) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Calcoliamo il passo.

Ricerca esatta

Annullando la derivata della $\phi(\alpha)$ nel punto $x^{(0)}$ si ottiene:

$$\frac{df(x^{(0)} + \alpha d^{(0)})}{d\alpha} = 2\alpha - 2 = 0$$

da cui $\alpha^{(0)} = 1$. Il nuovo punto è quindi:

$$x^{(1)} = x^{(0)} + \alpha^{(0)} d^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Metodo di Armijo

Inizializziamo i parametri del metodo nel modo seguente: $\gamma = 10^{-2}$, $\sigma = 0,5$.

La prima volta che il metodo di Armijo è applicato, uno dei metodi per calcolare il valore dello spostamento iniziale (α_0), è quello di approssimare la $\phi(\alpha)$ con una funzione quadratica del tipo $a\alpha^2 + b\alpha + c$, passante per $(0, \phi(0))$, con derivata $\phi'(0)$ in $\alpha = 0$, e avente un valore minimo assegnato u^*, dove u^* è una stima del minimo della funzione $\phi(\alpha)$. Il valore α_0 è dato dal punto che minimizza la parabola. In generale, si ha quindi che:

$$c = \phi(0), b = \phi'(0), a = \frac{b^2}{4(c - u^*)}$$

da cui

$$\alpha_0 = -\frac{2(c - u^*)}{b}$$

Nel nostro caso la $\phi(\alpha)$ è giá una funzione quadratica con minimo in $\alpha = 1$. Inizializziamo, quindi, il metodo di Armijo ponendo $\alpha_0 = 1$ e controlliamo se la condizione di arresto del metodo è verificata:

$$f(x^{(0)} + \alpha d^{(0)}) = -1$$

$$f(x^{(0)}) + \gamma\alpha\nabla f(x^{(0)})^T d^{(0)} = -0,02$$

La condizione

$$f(x^{(0)} + \alpha d^{(0)}) \leq f(x^{(0)}) + \gamma\alpha\nabla f(x^{(0)})^T d^{(0)}$$

è quindi verificata. Come d'altra parte c'era da aspettarsi dato che la ricerca esatta aveva fornito un passo proprio pari a 1.
Ponendo $\alpha^{(0)} = \alpha_0 = 1$ si ottiene lo stesso punto $x^{(1)}$ calcolato con la ricerca esatta.
Iterazione 1
Controlliamo la condizione di arresto dell'algoritmo

$$\nabla f(x^{(1)}) = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \neq 0$$

e quindi $x^{(1)}$ non è un punto stazionario.
La nuova direzione di discesa è:

$$d^{(1)} = -\nabla f(x^{(1)}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Calcoliamo il passo $\alpha^{(1)}$.
Ricerca esatta
Annullando la derivata della $\phi(\alpha) = f(x^{(1)} + \alpha d^{(1)})$ si ottiene:

$$\frac{df(x^{(1)} + \alpha d^{(1)})}{d\alpha} = 10\alpha - 2 = 0$$

da cui $\alpha^{(1)} = 1/5$. Il nuovo punto è quindi:

$$x^{(2)} = x^{(1)} + \alpha^{(1)} d^{(1)} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1/5 \\ 1/5 \end{bmatrix} = \begin{bmatrix} -0,8 \\ 1,2 \end{bmatrix}$$

Metodo di Armijo
Inizializzazione: $\alpha_0 = \alpha^{(0)} \frac{\nabla f(x^{(0)})^T d^{(0)}}{\nabla f(x^{(1)}) d^{(1)}} = 1$, $\gamma = 10^{-2}$, $\sigma = 0,5$.
Poniamo $\alpha = \alpha_0$ e controlliamo se la condizione di arresto del metodo è verificata:

$$f(x^{(1)} + \alpha d^{(1)}) = 2$$

$$f(x^{(1)}) + \gamma\alpha\nabla f(x^{(1)})^T d^{(1)} = -1,02$$

La condizione di arresto non è verificata.
Decrementiamo il valore del passo: $\alpha = \sigma\alpha = 0,5$, e controlliamo nuovamente la condizione di arresto nel punto

$$x^{(1)} + \alpha d^{(1)} = \begin{bmatrix} -1/2 \\ 3/2 \end{bmatrix}$$

$$f(x^{(1)} + \alpha d^{(1)}) = -3/4$$

$$f(x^{(1)}) + \gamma\alpha\nabla f(x^{(1)})^T d^{(1)} < -1$$

La condizione di arresto non è verificata.
Decremento del passo: $\alpha = \sigma\alpha = 0,25$. Verifica della condizione di arresto nel punto

$$x^{(1)} + \alpha d^{(1)} = \begin{bmatrix} -3/4 \\ 5/4 \end{bmatrix}$$

$$f(x^{(1)} + \alpha d^{(1)}) = -1,187$$

$$f(x^{(1)}) + \gamma\alpha\nabla f(x^{(1)})^T d^{(1)} = -1,005$$

La condizione di arresto è verificata. Il metodo di Armijo restituisce quindi il passo $\alpha^{(1)} = 0,25$. Il nuovo punto è:

$$x^{(2)} = x^{(1)} + \alpha d^{(1)} = \begin{bmatrix} -3/4 \\ 5/4 \end{bmatrix} = \begin{bmatrix} -0,75 \\ 1,25 \end{bmatrix}$$

Iterazione 2
Controlliamo se i punti trovati all'Iterazione 1 con i due diversi metodi di ricerca del passo soddisfano le condizioni necessarie del primo ordine:

$$\nabla f(-0,8; 1,2) = \begin{bmatrix} 1/5 \\ -1/5 \end{bmatrix} \neq 0$$

$$\nabla f(-0,75; 1,25) = \begin{bmatrix} 1/2 \\ 0 \end{bmatrix} \neq 0$$

Ricerca esatta

$$x^{(2)} = \begin{bmatrix} -4/5 \\ 6/5 \end{bmatrix} = \begin{bmatrix} -0,8 \\ 1,2 \end{bmatrix}$$

La nuova direzione di discesa è in questo caso:

$$d^{(2)} = -\nabla f(x^{(2)}) = \begin{bmatrix} -1/5 \\ 1/5 \end{bmatrix}$$

Calcoliamo il passo $\alpha^{(2)}$.
Annullando la derivata della $\phi(\alpha) = f(x^{(2)} + \alpha d^{(2)})$ si ottiene:

$$\frac{df(x^{(2)} + \alpha d^{(2)})}{d\alpha} = \frac{2\alpha - 2}{25} = 0$$

da cui $\alpha^{(2)} = 1$. Il nuovo punto è quindi:

$$x^{(3)} = x^{(2)} + \alpha^{(2)} d^{(2)} = \begin{bmatrix} -1 \\ 7/5 \end{bmatrix} = \begin{bmatrix} -1 \\ 1,4 \end{bmatrix}$$

Metodo di Armijo

$$x^{(2)} = \begin{bmatrix} -3/4 \\ 5/4 \end{bmatrix} = \begin{bmatrix} -0,75 \\ 1,25 \end{bmatrix}$$

La nuova direzione di discesa è in questo caso:

$$d^{(2)} = -\nabla f(x^{(2)}) = \begin{bmatrix} -1/2 \\ 0 \end{bmatrix}$$

Inizializzazione: $\alpha_0 = \alpha^{(1)} \frac{\nabla f(x^{(1)})^T d^{(1)}}{\nabla f(x^{(2)})^T d^{(2)}} = 2$, $\gamma = 10^{-2}$, $\sigma = 0,5$.
Applicando Armijo, il metodo si arresta dopo 4 iterazioni trovando $\alpha^{(2)} = 0,25$, da cui il punto

$$x^{(3)} = \begin{bmatrix} -7/8 \\ 5/4 \end{bmatrix} = \begin{bmatrix} -0,875 \\ 1,25 \end{bmatrix}$$

Si può verificare che entrambi i punti trovati non sono stazionari. Si deve dunque proseguire. Non terminiamo l'algoritmo. In Figura 3 sono riportate le sequenze di punti trovate. Notare che il metodo del gradiente con ricerca esatta si muove lungo direzioni successive che sono ortogonali. In Figura 4 è riportato il listato in *Matlab* del metodo del gradiente con ricerca del passo di Armijo per la funzione in esame.

Metodo di Newton

Applichiamo il metodo di Newton. Si noti che, poiché la funzione è quadratica e convessa, il metodo di Newton "puro" converge in un solo passo.
La direzione di Newton è

$$d^{(k)} = -(\nabla^2 f(x^{(k)}))^{-1} \nabla f(x^{(k)})$$

Si ha

$$(\nabla^2 f(x))^{-1} = \frac{1}{4} \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}$$

che è sempre definita positiva.

Iterazione 1

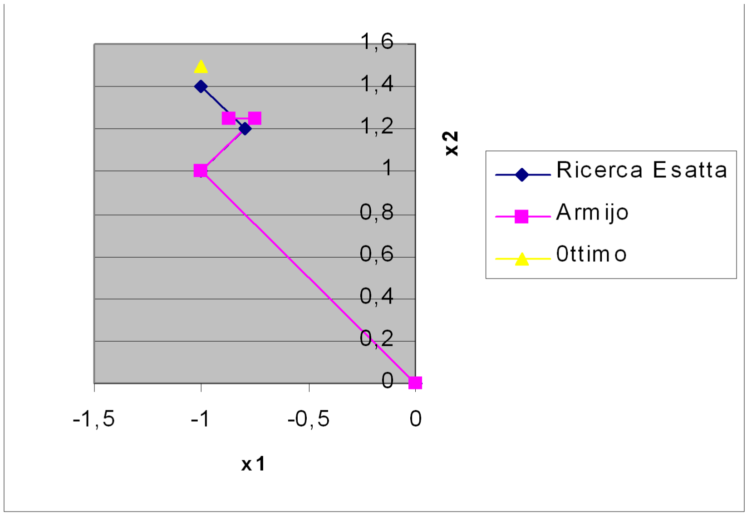

Figure 3: Metodo del gradiente con ricerca esatta e Armijo.

poniamo

$$d^{(0)} = -(\nabla^2 f(x^{(0)}))^{-1}\nabla f(x^{(0)}) = -\frac{1}{4}\begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 3/2 \end{bmatrix}$$

e $\alpha^{(0)} = 1$. Come era da aspettarsi, il nuovo punto

$$x^{(1)} = x^{(0)} + \alpha^{(0)} d^{(0)} = \begin{bmatrix} -1 \\ 3/2 \end{bmatrix}$$

è la soluzione ottima del problema.

Esempio 12 *Si consideri la funzione di Rosembrock dell'Esempio 5*

$$f(x_1, x_2) = (x_2 - x_1^2)^2 + (1 - x_1)^2$$

in cui si è posto $c = 1$. Applicare il metodo di Newton a partire dal punto

$$x^{(0)} = \begin{bmatrix} 0 \\ 1/2 \end{bmatrix}$$

Il gradiente e la matrice Hessiana della f sono rispettivamente:

$$\nabla f = \begin{bmatrix} -4x_1(x_2 - x_1^2) - 2(1 - x_1) \\ 2(x_2 - x_1^2) \end{bmatrix}$$

$$\nabla^2 f = 2\begin{bmatrix} 2(3x_1^2 - x_2) + 1 & -2x_1 \\ -2x_1 & 1 \end{bmatrix}$$

Nel punto iniziale $x^{(0)}$ si ha:

$$\nabla f(x^{(0)}) = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

Metodo del gradiente

```
%Punto iniziale
x(1)=0;
x(2)=0;
%funzione da minimizzare
f= x(1)-x(2)+2*x(1)^2+x(2)^2+2*x(1)*x(2);
%gradiente della funzione f
G = [(4*x(1)+2*x(2)+1) (2*x(2)+2*x(1)-1)];
k=0;
M=100;
while (k<M)
    if ((G < 0.0000001)&(G >= 0.0))
        break
    end
    %calcolo della direzione
    d=-G;
    %metodo di Armijo
    %Calcolo di a_0
    if (k==0)
        a_0=1
    else
        a_0=a_old *(G_old*d_old')/(G*d');
    end
    %inizializzazione Armijo
    a_old=a;
    a=a_0; gamma=0.001; sigma=0.5;
    x1=x+a*d;
    f1= x1(1)-x1(2)+2*x1(1)^2+x1(2)^2+2*x1(1)*x1(2);
    while (f1> (f+(gamma*a*(G*d'))))
        a=sigma*a;
        x1=x+a*d;
        f1= x1(1)-x1(2)+2*x1(1)^2+x1(2)^2+2*x1(1)*x1(2);
    end
    %fine Armijo
    %calcolo del nuovo punto e del valore della funzione
    x=x+a*d;
    f= x(1)-x(2)+2*x(1)^2+x(2)^2+2*x(1)*x(2);
    %gradiente e direzione all'iterazione corrente
    G_old=G;
    d_old=d;
    %nuovo Gradiente
    G = [(4*x(1)+2*x(2)+1) (2*x(2)+2*x(1)-1)];
    k=k+1;
end
```

Figure 4: Implementazione in Matlab del metodo del gradiente.

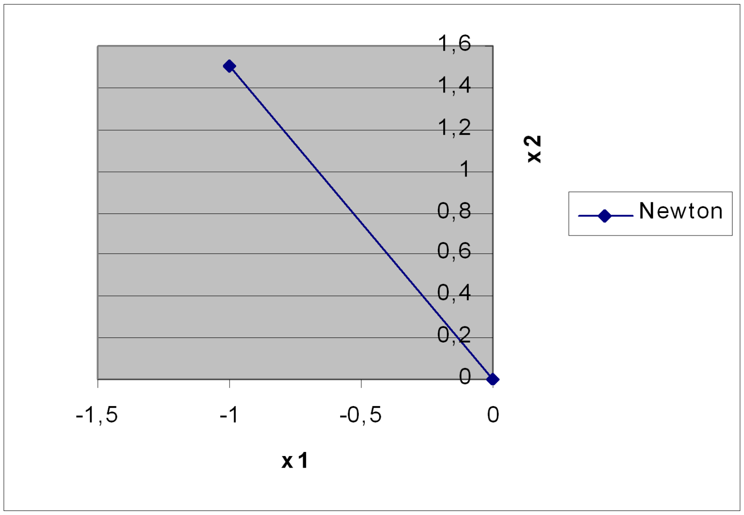

Figure 5: Metodo di Newton.

$$\nabla^2 f(x^{(0)}) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Notare che poichè $|\nabla^2 f(x^{(0)})| = 0$ non può essere utilizzato il metodo di Newton nella versione "pura". Scegliamo quindi come direzione di discesa la direzione dell'antigradiente:

$$d^{(0)} = -\nabla f(x^{(0)}) = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

Applichiamo Armijo per la ricerca del passo a partire dal valore $\alpha = \alpha_0 = 1$.
Inizializziamo i parametri del metodo nel modo seguente: $\gamma = 10^{-3}$, $\sigma = 0,5$.
Controlliamo se la condizione di arresto del metodo di Armijo è verificata:

$$x^{(0)} + \alpha d^{(0)} = \begin{bmatrix} 0 \\ 1/2 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ -1/2 \end{bmatrix}$$

$$f(x^{(0)} + \alpha d^{(0)}) = 7,25$$

$$f(x^{(0)}) + \gamma\alpha\nabla f(x^{(0)})^T d^{(0)} = 1,245$$

La condizione

$$f(x^{(0)} + \alpha d^{(0)}) \leq f(x^{(0)}) + \gamma\alpha\nabla f(x^{(0)})^T d^{(0)}$$

non è quindi verificata.
Poniamo $\alpha = \sigma\alpha = 1/2$. Controlliamo la condizione di arresto:

$$x^{(0)} + \alpha d^{(0)} = \begin{bmatrix} 0 \\ 1/2 \end{bmatrix} + \begin{bmatrix} 1 \\ -1/2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$f(x^{(0)} + \alpha d^{(0)}) = 1$$

$$f(x^{(0)}) + \gamma\alpha\nabla f(x^{(0)})^T d^{(0)} = 1,2475$$

Il passo $\alpha^{(0)}$ è quindi 1/2.
Il nuovo punto è

$$x^{(1)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

e si ha

$$\nabla f(x^{(1)}) = \begin{bmatrix} 4 \\ -2 \end{bmatrix} \neq 0$$

$$\nabla^2 f(x^{(1)}) = \begin{bmatrix} 14 & -4 \\ -4 & 2 \end{bmatrix}$$

$x^{(1)}$ non è un punto stazionario, si deve dunque proseguire. Non terminiamo l'esercizio, si noti che poichè ora $|\nabla^2 f(x^{(1)})| \neq 0$ possiamo utilizzare la direzione di Newton per il calcolo del punto $x^{(2)}$.

Esempio 13 *Data la funzione $f(x) = 2x_1^2 + x_2^2 - x_1x_2 - x_1$, trovare un punto di minimo analiticamente e poi cercarlo applicando il metodo del gradiente con line search esatta a partire dal punto* $(1, 1)$.

Il gradiente della funzione è $\nabla f(x) = (4x_1 - x_2 - 1, 2x_2 - x_1)^T$. Imponendo le condizioni del I ordine, abbiamo il sistema

$$\begin{cases} 4x_1 - x_2 - 1 = 0 \\ 2x_2 - x_1 = 0 \end{cases}$$

da cui otteniamo il punto $(\frac{2}{7}, \frac{1}{7})$. La matrice hessiana è data da

$$\nabla^2 f(x) = [\begin{array}{cc} 4 & -1 \\ -1 & 2 \end{array}]$$

ed è definita positiva ovunque, per cui il punto $(\frac{2}{7}, \frac{1}{7})$ è un punto di minimo locale. D'altra parte, si può osservare che la funzione è quadratica, ed essendo l'hessiana sempre definita positiva, essa è anche convessa. Si conclude quindi che il punto $(\frac{2}{7}, \frac{1}{7})$ è un punto di minimo globale.

Proviamo ora a raggiungere il punto di minimo a partire dal punto $(1, 1)$ applicando il metodo del gradiente con line search esatta. Poichè il gradiente della f nel punto $(1, 1)$ è $\nabla f(1, 1) = (2, 1)^T$, la direzione di maggior discesa in $(1, 1)$ è data da $d = -\nabla f(1, 1) = (-2, -1)^T$.

Per la ricerca del passo, facendo una line search esatta, esprimiamo la $\phi(\alpha)$ come

$$\phi(\alpha) = f(x_0 + \alpha d_0) = f(\begin{array}{c} 1-2\alpha \\ 1-\alpha \end{array}) =$$

$$2(1-2\alpha)^2 + (1-\alpha)^2 - (1-2\alpha)(1-\alpha) - (1-2\alpha) = 7\alpha^2 - 5\alpha + 1$$

da cui, imponendo $\phi'(\alpha) = 0$, si ottiene $\alpha = 5/14$. Il nuovo punto è allora $x_1 = (\frac{2}{7}, \frac{9}{14})$. In x_1 non è soddisfatta la condizione necessaria di ottimalità del primo ordine, essendo $\nabla f(x_1) = (1,1)^T$. Dunque, il metodo procede allo stesso modo, con la ricerca di un nuovo punto x_2 a partire dal punto x_1.

Esempio 14 *Trovare un punto di minimo della funzione dell'esercizio precedente applicando il metodo del gradiente con il metodo di Armijo a partire dal punto $x_0 = (1,1)$ $[\alpha_0 = 1, \gamma = 10^{-2}, \sigma = 0.5]$.*

Come nell'esercizio precedente, abbiamo $d_0 = -\nabla f(1,1) = (-2,-1)^T$. Stavolta, tuttavia, il passo è da determinarsi con il metodo di Armijo. Come richiesto dall'esercizio, inizializziamo Armijo ponendo $\alpha_0 = 1$ e vediamo se è soddisfatta la condizione di sufficiente riduzione. In caso affermativo, fisseremo $\alpha = \alpha_0$ e procederemo verso il nuovo punto, altrimenti, porremo $\alpha_1 = \sigma\alpha_0$ e ripeteremo il test. Il metodo di Armijo infatti prevede di ridurre progressivamente il valore di α lungo la direzione dell'antigradiente fino a soddisfare la condizione di sufficiente riduzione. Abbiamo:

$$f(x_0 + \alpha_0 d_0) = f((1,1)^T + (-2,-1)^T) = f(-1,0) = 3$$

$$f(x_0) + \gamma\alpha_0 \nabla f(x_0)^T d_0 = f(1,1) + 10^{-2}(2,1)^T(-2,-1) = 0,95$$

la condizione non è verificata, per cui poniamo $\alpha_1 = \sigma\alpha_0 = 1/2$ e ripetiamo il test:

$$f(x_0 + \alpha_1 d_0) = f(0, 1/2) = 1/4$$

$$f(x_0) + \gamma\alpha_1 \nabla f(x_0)^T d_0 = f(1,1) + \frac{10^{-2}(2,1)^T(-2,-1)}{2} = 0,975$$

Stavolta la condizione è verificata (dato che $0,25 < 0,975$), per cui prendiamo all'iterazione 0 il passo $\alpha = 1/2$, e il nuovo punto è $x_1 = (0, 1/2)^T$. Il gradiente della funzione nel nuovo punto è $\nabla f(x_1) = (-3/2, 1)^T \neq (0,0)^T$, per cui il metodo del gradiente procede come sopra cercando un nuovo punto x_2 a partire dal punto x_1.

Esempio 15 *Considerando la funzione dell'esercizio 5, trovare il punto di minimo applicando il metodo di Newton a partire dal punto $x_0 = (1,1)$.*

A partire dal punto dato, la direzione di Newton è data da:

$$d_0 = -(\nabla^2 f(x_0))^{-1} \nabla f(x_0)$$

Risulta:

$$\nabla^2 f(x) = [\begin{array}{cc} 4 & -1 \\ -1 & 2 \end{array}]$$

Essa è ovunque definita positiva, per cui è sempre possibile scegliere la direzione di Newton come direzione di discesa. Risulta dunque

$$(\nabla^2 f(x_0))^{-1} = \frac{1}{7}[\begin{array}{cc} 2 & 1 \\ 1 & 4 \end{array}]$$

da cui

$$d_0 = (-\frac{5}{7} \quad -\frac{6}{7})^T$$

A titolo di prova, verifichiamo che si tratti di una direzione di discesa:

$$\nabla f(x_0)^T d_0 = (2, 1)^T (-\frac{5}{7}, -\frac{6}{7}) < 0$$

Il nuovo punto è allora

$$x_1 = x_0 + d_0 = (1, 1)^T + (-\frac{5}{7}, -\frac{6}{7})^T = (2/7, 1/7)^T$$

Come avevamo immaginato, essendo la funzione quadratica, il metodo converge al punto stazionario in un solo passo.