Assignment 4

Elisa Piolai Mat. 856591

Esercizio 1

Si consideri il seguente problema di ottimizzazione non lineare

$$\min f(x,y) = 4(x-3)^2 + 2xy + 4y^2.$$

Si applichino due iterazioni del metodo del gradiente effettuando la line-search in modo esatto, a partire dal punto $A = (5,-2)^T$.

1 metodo del gradiente

$$f(x,y) = 4(x-3)^2 + 2xy + 4y^2$$

Si procede a calcolare il gradiente della funzione.

$$\frac{f(x,y)}{\partial x} = 8x + 2y - 24$$

$$\frac{f(x,y)}{\partial y} = 2x + 8y$$

$$\boxed{\nabla f(x,y) = \begin{bmatrix} 8x+2y-24 \\ 2x+8y \end{bmatrix}}$$ gradiente

Dato che si tratta di un problema di minimo, ad una generica iterazione del metodo del gradiente si avrà:

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ y_k \end{pmatrix} - \alpha \begin{pmatrix} \frac{f(x_k,y_k)}{\partial x} \\ \frac{f(x_k,y_k)}{\partial y} \end{pmatrix}$$

(dove α indica di quanto ci si sposta e il gradiente negativo la direzione di decrescita)

Si calcola α nel modo seguente:

$$\alpha^* = \min_{\alpha \geq 0} f\left(\begin{pmatrix} x_k \\ y_k \end{pmatrix} - \alpha \begin{pmatrix} \frac{f(x_k,y_k)}{\partial x} \\ \frac{f(x_k,y_k)}{\partial y} \end{pmatrix} \right)$$

ITERAZIONE 1

Si considera il punto corrente A:

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \end{pmatrix}$$

$$\nabla f(x,y) = \begin{bmatrix} 8x+2y-24 \\ 2x+8y \end{bmatrix}$$

Si calcola il gradiente nel punto

$$\nabla f(5,-2) = \begin{bmatrix} 12 \\ -6 \end{bmatrix}$$

Il nuovo punto sara' quindi:

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} - \alpha \begin{pmatrix} \frac{f(x_0,y_0)}{\partial x} \\ \frac{f(x_0,y_0)}{\partial y} \end{pmatrix} \longrightarrow \begin{pmatrix} 5 \\ -2 \end{pmatrix} - \alpha \begin{pmatrix} 12 \\ -6 \end{pmatrix} = \begin{pmatrix} 5-12\alpha \\ -2+6\alpha \end{pmatrix}$$

$$\alpha^* = \min_{\alpha \geq 0} f(5-12\alpha, -2+6\alpha)$$

Dato che la funzione e':

$$f(x,y) = 4(x-3)^2 + 2xy + 4y^2$$

$$\begin{aligned} f(5-12\alpha, -2+6\alpha) &= 4(5-12\alpha-3)^2 + 2(5-12\alpha)(-2+6\alpha) + 4(-2+6\alpha)^2 \\ &= 4(2-12\alpha)^2 + 2(-10+30\alpha+24\alpha-72\alpha^2) + 4(36\alpha^2+4-24\alpha) \\ &= 4(4+144\alpha^2-48\alpha) - 20 + 60\alpha + 48\alpha - \cancel{144\alpha^2} + \cancel{144\alpha^2} + 16 - 96\alpha \\ &= 16 + 576\alpha^2 - 192\alpha - 20 + 60\alpha + 48\alpha + 16 - 96\alpha \\ &= 576\alpha^2 - 180\alpha + 12 \end{aligned}$$

$$\alpha^\circ = \min_{\alpha \geq 0} g(\alpha) \qquad g(\alpha) = 576\alpha^2 - 180\alpha + 12$$

$$g'(\alpha) = 1152\alpha - 180 = 0 \longrightarrow \alpha^\circ = \frac{180}{1152} = \frac{5}{32} \longrightarrow \text{punto di minimo}$$

$$\begin{pmatrix} 5-12\alpha \\ -2+6\alpha \end{pmatrix} = \begin{pmatrix} 5-12\cdot\frac{5}{32} \\ -2+6\cdot\frac{5}{32} \end{pmatrix} = \begin{pmatrix} \frac{25}{8} \\ -\frac{17}{16} \end{pmatrix} \longrightarrow \text{nuovo punto trovato}$$

ITERAZIONE 2

Si considera il punto corrente:

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} \frac{25}{8} \\ -\frac{17}{16} \end{pmatrix}$$

$$\nabla f(x,y) = \begin{bmatrix} 8x+2y-24 \\ 2x+8y \end{bmatrix}$$

Si calcola il gradiente nel punto

$$\nabla f\left(\frac{25}{8}, -\frac{17}{16}\right) = \begin{bmatrix} -\frac{9}{8} \\ -\frac{9}{4} \end{bmatrix}$$

Il nuovo punto sara' quindi:

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} - \alpha \begin{pmatrix} \frac{f(x_1,y_1)}{\partial x} \\ \frac{f(x_1,y_1)}{\partial y} \end{pmatrix} \longrightarrow \begin{pmatrix} \frac{25}{8} \\ -\frac{17}{16} \end{pmatrix} - \alpha \begin{pmatrix} -\frac{9}{8} \\ -\frac{9}{4} \end{pmatrix} = \begin{pmatrix} \frac{9}{8}\alpha + \frac{25}{8} \\ \frac{9}{4}\alpha - \frac{17}{16} \end{pmatrix}$$

$$\alpha^* = \min_{\alpha \geq 0} f\left(\frac{9}{8}\alpha + \frac{25}{8}, \frac{9}{4}\alpha - \frac{17}{16}\right)$$

Dato che la funzione e':

$$f(x,y) = 4(x-3)^2 + 2xy + 4y^2$$

$$\begin{aligned}
f\left(\frac{9}{8}\alpha + \frac{25}{8}, \frac{9}{4}\alpha - \frac{17}{16}\right) &= 4\left(\frac{9}{8}\alpha + \frac{25}{8} - 3\right)^2 + 2\left(\frac{9}{8}\alpha + \frac{25}{8}\right)\left(\frac{9}{4}\alpha - \frac{17}{16}\right) + 4\left(\frac{9}{4}\alpha - \frac{17}{16}\right)^2 \\
&= 4\left(\frac{9}{8}\alpha + \frac{1}{2}\right)^2 + \frac{81}{16}\alpha^2 - \frac{153}{64}\alpha + \frac{225}{16}\alpha - \frac{425}{64} + 4\left(\frac{81}{16}\alpha^2 + \frac{289}{256} - \frac{153}{32}\alpha\right) \\
&= 4\left(\frac{81}{64}\alpha^2 + \frac{1}{64} + \frac{9}{32}\alpha\right) + \frac{81}{16}\alpha^2 - \frac{153}{64}\alpha + \frac{225}{16}\alpha - \frac{425}{64} + \frac{81}{4}\alpha^2 + \frac{289}{64} - \frac{153}{8}\alpha \\
&= \frac{81}{16}\alpha^2 + \frac{1}{16} + \frac{9}{8}\alpha + \frac{81}{16}\alpha^2 - \frac{153}{64}\alpha + \frac{225}{16}\alpha - \frac{425}{64} + \frac{81}{4}\alpha^2 + \frac{289}{64} - \frac{153}{8}\alpha \\
&= \frac{81+81+324}{16}\alpha^2 + \frac{72-153+900-1224}{64}\alpha + \frac{4-425+289}{64} \\
&= \frac{\cancel{486}^{243}}{\cancel{16}_{8}}\alpha^2 - \frac{405}{64}\alpha - \frac{\cancel{132}^{33}}{\cancel{64}_{16}} = \frac{243}{8}\alpha^2 - \frac{405}{64}\alpha - \frac{33}{16}
\end{aligned}$$

$$\alpha' = \min_{\alpha \geq 0} g(\alpha) \qquad g(\alpha) = \frac{243}{8}\alpha^2 - \frac{405}{64}\alpha - \frac{33}{16}$$

$$g'(\alpha) = \frac{243}{4}\alpha - \frac{405}{65} \longrightarrow \alpha' = \frac{5}{48}$$ → punto di minimo

$$\begin{pmatrix} \frac{9}{8}\alpha + \frac{25}{8} \\ \frac{9}{4}\alpha - \frac{17}{16} \end{pmatrix} = \begin{pmatrix} \frac{9}{8}\cdot\frac{5}{48} + \frac{25}{8} \\ \frac{9}{4}\cdot\frac{5}{48} - \frac{17}{16} \end{pmatrix} = \begin{pmatrix} \frac{415}{128} \\ -\frac{53}{64} \end{pmatrix}$$ → nuovo punto trovato

Esercizio 2

Si consideri il seguente problema di ottimizzazione non lineare

$$\min f(x,y) = 4(x-3)^2 + 2xy + 4y^2.$$

Si applichi il metodo di Newton a partire dal punto $A = (5,-2)^T$.

Commentare la differenza, in termini di convergenza, rispetto al metodo del gradiente.

2 metodo di newton

$$f(x,y) = 4(x-3)^2 + 2xy + 4y^2$$

Si procede a calcolare il gradiente della funzione.

$$\frac{f(x,y)}{\partial x} = 8x + 2y - 24$$

$$\frac{f(x,y)}{\partial y} = 2x + 8y$$

$$\boxed{\nabla f(x,y) = \begin{bmatrix} 8x + 2y - 24 \\ 2x + 8y \end{bmatrix}}$$ gradiente

Si calcola quindi la matrice Hessiana

$$H_f(x,y) = \begin{bmatrix} 8 & 2 \\ 2 & 8 \end{bmatrix}$$

Ad una generica iterazione del metodo di Newton si avrà:

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ y_k \end{pmatrix} - H_f(x,y)^{-1} \nabla f(x_k y_k)$$

Si calcola il gradiente nel punto

$$\nabla f(5,-2) = \begin{bmatrix} 12 \\ -6 \end{bmatrix}$$

Si deve ora calcolare

$$H_f(x,y)^{-1} \nabla f(x_k y_k) \longrightarrow \begin{bmatrix} 8 & 2 \\ 2 & 8 \end{bmatrix}^{-1} \begin{bmatrix} 12 \\ -6 \end{bmatrix}$$

$$-H_f(x,y)^{-1} \nabla f(x_k, y_k) = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

$$H_f(x,y) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = -\nabla f(x_k, y_k)$$

$$-\begin{bmatrix} 8 & 2 \\ 2 & 8 \end{bmatrix}^{-1} \begin{bmatrix} 12 \\ -6 \end{bmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \longrightarrow \begin{bmatrix} 8 & 2 \\ 2 & 8 \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{bmatrix} -12 \\ 6 \end{bmatrix}$$

$$\begin{cases} 8v_1 + 2v_2 = -12 \\ 2v_1 + 8v_2 = 6 \end{cases} \Rightarrow \begin{cases} v_1 = -\frac{9}{5} \\ v_2 = \frac{6}{5} \end{cases}$$

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \end{pmatrix} + \begin{pmatrix} -9/5 \\ 6/5 \end{pmatrix} = \begin{pmatrix} 16/5 \\ -4/5 \end{pmatrix}$$

Si verifica quindi che il punto trovato sia il punto di minimo.

Si calcola il valore del gradiente

$\nabla f\left(\frac{16}{5}, -\frac{4}{5}\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \longrightarrow$ il punto e' un candidato ad essere un punto di minimo

Si procede calcolando gli autovalori della matrice Hessiana:

$$H_f(x,y) = \begin{bmatrix} 8 & 2 \\ 2 & 8 \end{bmatrix} \longrightarrow \begin{bmatrix} 8-\lambda & 2 \\ 2 & 8-\lambda \end{bmatrix}$$

$$(8-\lambda)(8-\lambda) - 4 = 0$$

$$64 + \lambda^2 - 16\lambda - 4 = 0$$

$$\lambda^2 - 16\lambda + 60 = 0$$

$$\lambda_{1/2} = 8 \pm \sqrt{64-60} = 8 \pm \sqrt{4} = 8 \pm 2$$

$$\lambda_1 = 10$$

$$\lambda_2 = 6$$

Entrambi gli autovalori sono positivi $\longrightarrow$ la matrice e' definita positiva

$\Rightarrow$ il punto $\begin{pmatrix} 16/5 \\ -4/5 \end{pmatrix}$ e' il minimo della funzione

Rispetto al metodo del gradiente la funzione converge in una sola iterazione in quanto e' una funzione quatratica

si può inoltre far vedere che con il metodo del gradiente la funzione non converge alla seconda iterazione. Calcolando il gradiente, si osserva che questo non si annulla

$$\nabla f(x,y) = \begin{bmatrix} 8x + 2y - 24 \\ 2x + 8y \end{bmatrix} = \begin{bmatrix} 8 \cdot \frac{418}{128} - 2 \cdot \frac{53}{64} - 24 \\ 2 \cdot \frac{418}{28} - 8 \frac{53}{64} \end{bmatrix} = \begin{bmatrix} \frac{15}{32} \\ -\frac{3}{32} \end{bmatrix} \neq 0$$

Esercizio 3

Si consideri la funzione obiettivo $f(x, y) = x$ soggetta ai seguenti vincoli:

$$x^2 + (y-1)^2 \leq 4$$

$$x^2 + (y+1)^2 \leq 4$$

Si trovino i punti di massimo e di minimo della funzione utilizzando le condizioni KKT.

3 *condizioni KKT*

Relativamente a questo caso, abbiamo solo vincoli di disuguaglianza

Si calcolano i gradienti della funzione obiettivo e dei vincoli:

a. $f(x,y) = x$

$$\frac{\partial f(x,y)}{\partial x} = 1$$

$$\frac{\partial f(x,y)}{\partial y} = 0$$

b. $h_1(x,y) = x^2 + (y-1)^2 - 4 \leq 0$

$$\frac{\partial h_1(x,y)}{\partial x} = 2x$$

$$\frac{\partial h_1(x,y)}{\partial y} = 2y - 2$$

c. $h_2(x,y) = x^2 + (y+1)^2 - 4 \leq 0$

$$\frac{\partial h_2(x,y)}{\partial x} = 2x$$

$$\frac{\partial h_2(x,y)}{\partial y} = 2y + 2$$

Condizioni KKT nel dettaglio:

Condizione sulla derivata parziale rispetto a x

$$\frac{\partial f(x,y)}{\partial x} = -\mu_1 \frac{\partial h_1(x,y)}{\partial x} - \mu_2 \frac{\partial h_2(x,y)}{\partial x}$$

$$1 = -\mu_1 \cdot 2x - \mu_2 \cdot 2x$$

Condizione sulla derivata parziale rispetto a y

$$\frac{\partial f(x,y)}{\partial y} = -\mu_1 \frac{\partial h_1(x,y)}{\partial y} - \mu_2 \frac{\partial h_2(x,y)}{\partial y}$$

$$0 = -\mu_1 (2y-2) - \mu_2 (2y+2)$$

Condizioni di complementarieta'

1. $\mu_1 \cdot h_1(x,y) = 0$

   $\mu_1 \cdot (x^2 + (y-1)^2 - 4) = 0$

2. $\mu_2 \cdot h_2(x,y) = 0$

   $\mu_2 \cdot (x^2 + (y+1)^2 - 4) = 0$

Condizioni di ammissibilita' dei vincoli

1. $h_1(x,y) \leq 0$

   $x^2 + (y-1)^2 - 4 \leq 0$

2. $h_2(x,y) \leq 0$

   $x^2 + (y+1)^2 - 4 \leq 0$

Condizioni di non negativita'

$\mu_1, \mu_2, \mu_3 \geq 0$

Il sistema risultante sara' quindi:

punti di minimo

$$\begin{cases} 1 = -\mu_1 \cdot 2x - \mu_2 \cdot 2x \\ 0 = -\mu_1(2y-2) - \mu_2(2y+2) \\ \mu_1 \cdot (x^2 + (y-1)^2 - 4) = 0 \\ \mu_2 \cdot (x^2 + (y+1)^2 - 4) = 0 \\ x^2 + (y-1)^2 - 4 \leq 0 \\ x^2 + (y+1)^2 - 4 \leq 0 \\ \mu_1, \mu_2 \geq 0 \end{cases}$$

punti di massimo

$$\begin{cases} 1 = \mu_1 \cdot 2x + \mu_2 \cdot 2x \\ 0 = \mu_1(2y-2) + \mu_2(2y+2) \\ \mu_1 \cdot (x^2 + (y-1)^2 - 4) = 0 \\ \mu_2 \cdot (x^2 + (y+1)^2 - 4) = 0 \\ x^2 + (y-1)^2 - 4 \leq 0 \\ x^2 + (y+1)^2 - 4 \leq 0 \\ \mu_1, \mu_2 \geq 0 \end{cases}$$

Per risolvere il sistema si deve analizzare l'attivazione o meno dei due vincoli di disuguaglianza (4 casi in totale)

1. $\mu_1 = 0, \ \mu_2 = 0$
2. $\mu_1 > 0, \ \mu_2 = 0$
3. $\mu_1 = 0, \ \mu_2 > 0$
4. $\mu_1 > 0, \ \mu_2 > 0$

PUNTI DI MINIMO

1° caso

$$\begin{cases} 1 = 0 \\ 1 = 0 \\ \dots \end{cases}$$

impossibile

2° caso

$$\begin{cases} 0 = -\mu_1(2y-2) \rightarrow 2y-2=0 \rightarrow y=1 \\ x^2 + (y-1)^2 - 4 = 0 \rightarrow x^2 + y^2 - 2y - 3 = 0 \rightarrow x^2 + 1 - 2 - 3 = 0 \rightarrow x^2 = 4 \rightarrow x = \pm 2 \\ x^2 + (y+1)^2 - 4 \leq 0 \rightarrow \cancel{4} + (2)^2 - \cancel{4} \leq 0 \rightarrow 4 \leq 0 \\ \dots \end{cases}$$

impossibile

3° caso

$$\begin{cases} 0 = -\mu_2(2y+2) \rightarrow y=-1 \\ x^2+(y+1)^2-4=0 \rightarrow x^2+0-4=0 \rightarrow x=\pm 2 \\ x^2+(y-1)^2-4 \leq 0 \rightarrow 4+\cancel{4}-\cancel{4} \leq 0 \rightarrow 4 \leq 0 \\ \dots \end{cases}$$

impossibile

4° caso

$$\begin{cases} x^2+(y+1)^2-4=0 \rightarrow x^2=3-y^2-2y \rightarrow x^2=3 \rightarrow x=\pm\sqrt{3} \\ x^2+(y-1)^2-4=0 \rightarrow 3-\cancel{y^2}-2y+\cancel{y^2}+1-2y-4=0 \rightarrow y=0 \\ x^2+(y-1)^2-4 \leq 0 \rightarrow 3+1-4 \leq 0 \rightarrow 0 \leq 0 \\ x^2+(y+1)^2-4 \leq 0 \rightarrow 3+1-4 \leq 0 \rightarrow 0 \leq 0 \\ \dots \end{cases}$$

$(-\sqrt{3}, 0), (\sqrt{3}, 0)$ punti candidati ad essere minimo locale

$f(\sqrt{3}, 0) = \sqrt{3}$

$f(-\sqrt{3}, 0) = -\sqrt{3} \rightarrow$ valore più basso

$\boxed{A = (-\sqrt{3}, 0) \Rightarrow \text{punto di minimo globale}}$

PUNTI DI MASSIMO

1° caso

$$\begin{cases} 1 = 0 \\ 1 = 0 \\ \dots \end{cases}$$

impossibile

2° caso

$$\begin{cases} 0 = \mu_1(2y-2) \rightarrow 2y-2=0 \rightarrow y=1 \\ x^2+(y-1)^2-4=0 \rightarrow x^2+y^2-2y-3=0 \rightarrow x^2+1-2-3=0 \rightarrow x^2=4 \rightarrow x=\pm 2 \\ x^2+(y+1)^2-4 \leq 0 \rightarrow \cancel{4}+(2)^2-\cancel{4} \leq 0 \rightarrow 4 \leq 0 \\ \dots \end{cases}$$

impossibile

3° caso

$$\begin{cases} 0 = \mu_2(2y+2) \rightarrow y=-1 \\ x^2+(y+1)^2-4=0 \rightarrow x^2=4 \rightarrow x=\pm 2 \\ x^2+(y-1)^2-4 \leq 0 \rightarrow 4+\cancel{4}-\cancel{4} \leq 0 \rightarrow 4 \leq 0 \\ \dots \end{cases}$$

impossibile

4° caso

$$\begin{cases} x^2+(y+1)^2-4=0 \longrightarrow x^2=3-y^2-2y \longrightarrow x^2=3 \longrightarrow x=\pm\sqrt{3} \\ x^2+(y-1)^2-4=0 \longrightarrow 3-\cancel{y^2}-2y+\cancel{y^2}+1-2y-4=0 \longrightarrow y=0 \\ x^2+(y-1)^2-4\leq 0 \longrightarrow 3+1-4\leq 0 \longrightarrow 0\leq 0 \\ x^2+(y+1)^2-4\leq 0 \longrightarrow 3+1-4\leq 0 \longrightarrow 0\leq 0 \\ \dots \end{cases}$$

$(-\sqrt{3},0),(\sqrt{3},0)$ punti candidati ad essere massimo locale

$f(-\sqrt{3},0)=-\sqrt{3}$

$f(\sqrt{3},0)=\sqrt{3} \longrightarrow$ valore più alto

$\boxed{A=(\sqrt{3},0) \Rightarrow \text{punto di massimo globale}}$