Ottimizzazione dei Sistemi Complessi

G. Liuzzi[1]

Martedì 5 Aprile 2016

[1] Istituto di Analisi dei Sistemi ed Informatica IASI - CNR

Un (bel) passo indietro ... o avanti!

Consideriamo il problema *NON vincolato*

$$\begin{array}{ll} \min & f(x) \\ c.v. & x \in \mathbb{R}^n \end{array}$$

Supponiamo di conoscere una soluzione x^* del problema e siano verificate le **C.S. del II ordine** in x^*:

- $\nabla f(x^*) = 0$
- $\nabla^2 f(x^*)$ definita positiva

N.B. Attenzione! Se x^* è minimo locale stretto, non è detto che $\nabla^2 f(x^*)$ sia definita positiva!

Il metodo di Newton (1)

Vogliamo costruire una approsimazione del problema in un intorno di x^*

Sviluppo in serie di Taylor troncato al secondo ordine di $f(x)$:

$$
\begin{aligned}
q(x;x^*) &= f(x^*) + \nabla f(x^*)^\top (x - x^*) + \frac{1}{2}(x - x^*)^\top \nabla^2 f(x^*)(x - x^*) \\
f(x) &\simeq q(x;x^*) \\
q(x^*;x^*) &= f(x^*) \\
\bar{q}(d) &= q(x^* + d; x^*) = f(x^*) + \frac{1}{2} d^\top \nabla^2 f(x^*) d \\
\nabla \bar{q}(d) &= \nabla^2 f(x^*) d \\
\nabla^2 \bar{q}(d) &= \nabla^2 f(x^*)
\end{aligned}
$$

Il metodo di Newton (1)

Vogliamo costruire una approsimazione del problema in un intorno di x^*

Sviluppo in serie di Taylor troncato al secondo ordine di $f(x)$:

$$
\begin{aligned}
q(x;x^*) &= f(x^*) + \nabla f(x^*)^\top (x-x^*) + \frac{1}{2}(x-x^*)^\top \nabla^2 f(x^*)(x-x^*) \\
f(x) &\simeq q(x;x^*) \\
q(x^*;x^*) &= f(x^*) \\
\bar{q}(d) &= q(x^*+d;x^*) = f(x^*) + \frac{1}{2} d^\top \nabla^2 f(x^*) d \\
\nabla \bar{q}(d) &= \nabla^2 f(x^*) d \\
\nabla^2 \bar{q}(d) &= \nabla^2 f(x^*)
\end{aligned}
$$

Il metodo di Newton (1)

Vogliamo costruire una approsimazione del problema in un intorno di x^*

Sviluppo in serie di Taylor troncato al secondo ordine di $f(x)$:

$$
\begin{aligned}
q(x;x^*) &= f(x^*) + \nabla f(x^*)^\top (x - x^*) + \frac{1}{2}(x - x^*)^\top \nabla^2 f(x^*)(x - x^*) \\
f(x) &\simeq q(x;x^*) \\
q(x^*;x^*) &= f(x^*) \\
\bar{q}(d) &= q(x^* + d; x^*) = f(x^*) + \frac{1}{2} d^\top \nabla^2 f(x^*) d \\
\nabla \bar{q}(d) &= \nabla^2 f(x^*) d \\
\nabla^2 \bar{q}(d) &= \nabla^2 f(x^*)
\end{aligned}
$$

Il metodo di Newton (2)

$q(x;x^*)$ e quindi $\bar{q}(d)$ sono una "buona" approssimazione di f in un intorno di x^* quando:

- x^* è minimo locale (oltre che di f) anche per $q(x;x^*)$ e
- $d^* = 0$ è minimo locale di $\bar{q}(d)$.

Infatti, da

$$\begin{aligned} \nabla \bar{q}(d) &= \nabla^2 f(x^*) d = 0 \\ \nabla^2 \bar{q}(d) &= \nabla^2 f(x^*) \ d.p. \end{aligned}$$

segue che $d^* = 0$ è proprio l'**unico** minimo locale di $\bar{q}(d)$

Il metodo di Newton (2)

$q(x; x^*)$ e quindi $\bar{q}(d)$ sono una "buona" approssimazione di f in un intorno di x^* quando:

- x^* è minimo locale (oltre che di f) anche per $q(x; x^*)$ e
- $d^* = 0$ è minimo locale di $\bar{q}(d)$.

Infatti, da

$$\begin{aligned} \nabla \bar{q}(d) &= \nabla^2 f(x^*) d = 0 \\ \nabla^2 \bar{q}(d) &= \nabla^2 f(x^*) \ d.p. \end{aligned}$$

segue che $d^* = 0$ è proprio l'**unico** minimo locale di $\bar{q}(d)$

Il metodo di Newton (3)

Supponiamo ora di **non** conoscere x^* ma di disporre di una **stima** x_k di x^*

Idea

- definisco $q(x; x_k)$ e $\bar{q}_k(d)$
- minimizzo $q(x; x_k)$
- aggiorno la stima di x^* definendo x_{k+1}

Il metodo di Newton (3)

Supponiamo ora di **non** conoscere x^* ma

di disporre di una **stima** x_k di x^*

Idea

- definisco $q(x; x_k)$ e $\bar{q}_k(d)$
- minimizzo $q(x; x_k)$
- aggiorno la stima di x^* definendo x_{k+1}

Il metodo di Newton (4)

Naturalmente, risulta:

$$\begin{aligned}
\bar{q}_k(d) &= f(x_k) + \nabla f(x_k)^\top d + \frac{1}{2} d^\top \nabla^2 f(x_k) d \\
\nabla \bar{q}_k(d) &= \nabla f(x_k) + \nabla^2 f(x_k) d
\end{aligned}$$

$\nabla \bar{q}_k(d) = 0$ quando $\nabla f(x_k) + \nabla^2 f(x_k) d = 0$.
Se $\nabla^2 f(x_k)$ è non singolare, allora possiamo definire:

$$\begin{aligned}
d_k &= -\nabla^2 f(x_k)^{-1} \nabla f(x_k) && \textit{direzione di Newton} \\
x_{k+1} &= x_k + d_k && \textit{iterazione di Newton}
\end{aligned}$$

Il metodo di Newton (4)

Naturalmente, risulta:

$$\begin{aligned} \bar{q}_k(d) &= f(x_k) + \nabla f(x_k)^\top d + \frac{1}{2} d^\top \nabla^2 f(x_k) d \\ \nabla \bar{q}_k(d) &= \nabla f(x_k) + \nabla^2 f(x_k) d \end{aligned}$$

$\nabla \bar{q}_k(d) = 0$ quando $\nabla f(x_k) + \nabla^2 f(x_k) d = 0$.
Se $\nabla^2 f(x_k)$ è non singolare, allora possiamo definire:

$$\begin{aligned} d_k &= -\nabla^2 f(x_k)^{-1} \nabla f(x_k) && \textit{direzione di Newton} \\ x_{k+1} &= x_k + d_k && \textit{iterazione di Newton} \end{aligned}$$

Caso Vincolato (semplice)

Consideriamo il problema

$$\begin{array}{ll} \min_x & f(x) \\ c.v. & h(x) = 0 \end{array}$$

Supponiamo di conoscere una soluzione (x^*, μ^*) del problema e siano verificate le **C.S. del II ordine** in (x^*, μ^*):

- $\nabla_x L(x^*, \mu^*) = 0,\ h(x^*) = 0$
- $rg(\nabla h(x^*)) = p$ grad. dei vincoli di uguaglianza sono lin.indip.
- $d^\top \nabla_x^2 L(x^*, \mu^*) d > 0$ per ogni $d \neq 0$ tale che $\nabla h(x^*)^\top d = 0$

N.B. Attenzione! Se x^* è minimo locale stretto, certamente risulta (C.N. del II ordine)
$d^\top \nabla_x^2 L(x^*, \mu^*) d \geq 0$ per ogni $d \neq 0$ tale che $\nabla h(x^*)^\top d = 0$

Approssimazione

Vogliamo costruire una approssimazione (rispetto ad x, tenendo μ^* fisso) del problema vincolato in un intorno del punto x^*

Approssimazione – 1° tentativo

- Sviluppo in serie di Taylor troncato al II ordine di $f(x)$
- Sviluppo in serie di Taylor troncato al I ordine di $h(x)$

$$\begin{array}{ll} \min_x & f(x^*) + \nabla f(x^*)^\top (x - x^*) + \frac{1}{2}(x - x^*)^\top \nabla^2 f(x^*)(x - x^*) \\ c.v. & h(x^*) + \nabla h(x^*)^\top (x - x^*) = 0 \end{array}$$

$$\begin{array}{ll} \min_d & f(x^*) + \nabla f(x^*)^\top d + \frac{1}{2} d^\top \nabla^2 f(x^*) d \\ c.v. & h(x^*) + \nabla h(x^*)^\top d = 0 \end{array}$$

Vogliamo essere sicuri che $(x^*, \mu^*) = (d^* = 0, \mu^*)$ sia soluzione anche del problema approssimante

Approssimazione – 1° tentativo

- Sviluppo in serie di Taylor troncato al II ordine di $f(x)$
- Sviluppo in serie di Taylor troncato al I ordine di $h(x)$

$$
\begin{array}{ll}
\min_x & f(x^*) + \nabla f(x^*)^\top (x - x^*) + \frac{1}{2}(x - x^*)^\top \nabla^2 f(x^*)(x - x^*) \\
c.v. & h(x^*) + \nabla h(x^*)^\top (x - x^*) = 0
\end{array}
$$

$$
\begin{array}{ll}
\min_d & f(x^*) + \nabla f(x^*)^\top d + \frac{1}{2} d^\top \nabla^2 f(x^*) d \\
c.v. & h(x^*) + \nabla h(x^*)^\top d = 0
\end{array}
$$

Vogliamo essere sicuri che $(x^*, \mu^*) = (d^* = 0, \mu^*)$ sia soluzione anche del problema approssimante

Approssimazione – 1° tentativo

$$
\begin{aligned}
\nabla_d \bar{L} &= \nabla f(x^*) + \nabla^2 f(x^*) d + \nabla h(x^*) \bar{\mu} \\
\nabla_d^2 \bar{L} &= \nabla^2 f(x^*)
\end{aligned}
$$

Verifichiamo se $d = 0$, $\bar{\mu} = \mu^*$ è una coppia che soddisfa le C.N. di I e II ordine

Tuttavia, in generale non si verifica che

$$d^T \nabla^2 f(x^*) d \geq 0, \text{ per ogni } d \neq 0 : \nabla h(x^*)^T d = 0$$

cioè non valgono (in generale) le C.N. del II ordine e quindi (x^*, μ^*) non è soluzione del problema approssimante

Approssimazione – 1° tentativo

$$\begin{aligned} \nabla_d \bar{L} &= \nabla f(x^*) + \nabla^2 f(x^*) d + \nabla h(x^*) \bar{\mu} \\ \nabla_d^2 \bar{L} &= \nabla^2 f(x^*) \end{aligned}$$

Verifichiamo se $d = 0$, $\bar{\mu} = \mu^*$ è una coppia che soddisfa le C.N. di I e II ordine

$$\begin{aligned} \nabla f(x^*) + \nabla^2 f(x^*) d + \nabla h(x^*) \bar{\mu} &= \nabla_d \bar{L} \\ h(x^*) + \nabla h(x^*)^\top d &= 0 \end{aligned}$$

Tuttavia, in generale **non** si verifica che

$$d^\top \nabla^2 f(x^*) d \geq 0, \text{ per ogni } d \neq 0 : \ \nabla h(x^*)^\top d = 0$$

cioè **non** valgono (in generale) le C.N. del II ordine e quindi (x^*, μ^*) non è soluzione del problema approssimante

Approssimazione – 1° tentativo

$$\begin{aligned} \nabla_d \bar{L} &= \nabla f(x^*) + \nabla^2 f(x^*) d + \nabla h(x^*) \bar{\mu} \\ \nabla_d^2 \bar{L} &= \nabla^2 f(x^*) \end{aligned}$$

Verifichiamo se $d = 0$, $\bar{\mu} = \mu^*$ è una coppia che soddisfa le C.N. di I e II ordine

$$\begin{aligned} \nabla f(x^*) + \nabla h(x^*) \mu^* = \nabla_x L(x^*, \mu^*) &= 0 \\ h(x^*) &= 0 \end{aligned}$$

Tuttavia, in generale **non** si verifica che

$$d^\top \nabla^2 f(x^*) d \geq 0, \text{ per ogni } d \neq 0 : \ \nabla h(x^*)^\top d = 0$$

cioè **non** valgono (in generale) le C.N. del II ordine e quindi (x^*, μ^*) non è soluzione del problema approssimante

Approssimazione – 1° tentativo

$$
\begin{array}{rcl}
\nabla_d \bar{L} & = & \nabla f(x^*) + \nabla^2 f(x^*) d + \nabla h(x^*) \bar{\mu} \\
\nabla_d^2 \bar{L} & = & \nabla^2 f(x^*)
\end{array}
$$

Verifichiamo se $d = 0$, $\bar{\mu} = \mu^*$ è una coppia che soddisfa le C.N. di I e II ordine

$$
\begin{array}{rcl}
\nabla f(x^*) + \nabla h(x^*)\mu^* = \nabla_x L(x^*, \mu^*) & = & 0 \\
h(x^*) & = & 0
\end{array}
$$

Tuttavia, in generale **non** si verifica che

$$
d^\top \nabla^2 f(x^*) d \geq 0, \text{ per ogni } d \neq 0 : \ \nabla h(x^*)^\top d = 0
$$

cioè **non** valgono (in generale) le C.N. del II ordine e quindi (x^*, μ^*) non è soluzione del problema approssimante

Approssimazione – 2° tentativo

- Sviluppo in serie di Taylor troncato al II ordine di $L(x, \mu^*)$
- Sviluppo in serie di Taylor troncato al I ordine di $h(x)$

$$\begin{array}{ll} \min_x & L(x^*, \mu^*) + \nabla_x L(x^*, \mu^*)^\top (x - x^*) + \frac{1}{2}(x - x^*)^\top \nabla_x^2 L(x^*, \mu^*)(x - x^*) \\ c.v. & h(x^*) + \nabla h(x^*)^\top (x - x^*) = 0 \end{array}$$

$$\begin{array}{ll} \min_d & L(x^*, \mu^*) + \frac{1}{2} d^\top \nabla_x^2 L(x^*, \mu^*) d \\ c.v. & h(x^*) + \nabla h(x^*)^\top d = 0 \end{array}$$

Vogliamo essere sicuri che $(x^*, \mu^*) = (d^* = 0, \mu^*)$ sia soluzione anche del problema approssimante

Approssimazione – 2° tentativo

- Sviluppo in serie di Taylor troncato al II ordine di $L(x, \mu^*)$
- Sviluppo in serie di Taylor troncato al I ordine di $h(x)$

$$\begin{array}{ll} \min_x & L(x^*, \mu^*) + \nabla_x L(x^*, \mu^*)^\top (x - x^*) + \frac{1}{2}(x - x^*)^\top \nabla_x^2 L(x^*, \mu^*)(x - x^*) \\ c.v. & h(x^*) + \nabla h(x^*)^\top (x - x^*) = 0 \end{array}$$

$$\begin{array}{ll} \min_d & L(x^*, \mu^*) + \frac{1}{2} d^\top \nabla_x^2 L(x^*, \mu^*) d \\ c.v. & h(x^*) + \nabla h(x^*)^\top d = 0 \end{array}$$

Vogliamo essere sicuri che $(x^*, \mu^*) = (d^* = 0, \mu^*)$ sia soluzione anche del problema approssimante

Approssimazione – 2° tentativo

$$\begin{aligned}
\nabla_d \bar{L} &= \nabla_x^2 L(x^*, \mu^*) d + \nabla h(x^*) \bar{\mu} \\
\nabla_d^2 \bar{L} &= \nabla_x^2 L(x^*, \mu^*)
\end{aligned}$$

Verifichiamo se $d = 0$, $\bar{\mu} = \mu^*$ è una coppia che soddisfa le C.N. di I e II ordine

In generale non risulta $\nabla h(x^*)\mu^* = 0$. Si pensi a punti x^* "regolari" cioè tali per cui risultano lin. indipendenti le colonne di $\nabla h(x^*)$ e quindi $\nabla h(x^*)\mu = 0$ implica $\mu = 0 \neq \mu^*$.

Quindi, non valgono (in generale) le C.N. del I ordine, e quindi (x^*, μ^*) non è soluzione del problema approssimante.

Approssimazione – 2° tentativo

$$
\begin{aligned}
\nabla_d \bar{L} &= \nabla_x^2 L(x^*, \mu^*) d + \nabla h(x^*) \bar{\mu} \\
\nabla_d^2 \bar{L} &= \nabla_x^2 L(x^*, \mu^*)
\end{aligned}
$$

Verifichiamo se $d = 0$, $\bar{\mu} = \mu^*$ è una coppia che soddisfa le C.N. di I e II ordine

$$
\begin{aligned}
\nabla_x^2 L(x^*, \mu^*) d + \nabla h(x^*) \bar{\mu} &= \nabla_d \bar{L} \\
h(x^*) + \nabla h(x^*)^\top d &= 0
\end{aligned}
$$

In generale **non** risulta $\nabla h(x^*) \mu^* = 0$. Si pensi a punti x^* "regolari" cioè tali per cui risultano lin. indipendenti le colonne di $\nabla h(x^*)$ e quindi $\nabla h(x^*) \mu = 0$ implica $\mu = 0 \neq \mu^*$.

Quindi, **non** valgono (in generale) le C.N. del I ordine, e quindi (x^*, μ^*) non è soluzione del problema approssimante

Approssimazione – 2° tentativo

$$\begin{aligned}\nabla_d \bar{L} &= \nabla_x^2 L(x^*, \mu^*) d + \nabla h(x^*) \bar{\mu} \\ \nabla_d^2 \bar{L} &= \nabla_x^2 L(x^*, \mu^*)\end{aligned}$$

Verifichiamo se $d = 0$, $\bar{\mu} = \mu^*$ è una coppia che soddisfa le C.N. di I e II ordine

$$\begin{aligned}\nabla_d \bar{L}(0, \mu^*) &= \nabla h(x^*) \mu^* \\ h(x^*) &= 0\end{aligned}$$

In generale **non** risulta $\nabla h(x^*)\mu^* = 0$. Si pensi a punti x^* "regolari" cioè tali per cui risultano lin. indipendenti le colonne di $\nabla h(x^*)$ e quindi $\nabla h(x^*)\mu = 0$ implica $\mu = 0 \neq \mu^*$.

Quindi, **non** valgono (in generale) le C.N. del I ordine, e quindi (x^*, μ^*) non è soluzione del problema approssimante

Approssimazione – 2° tentativo

$$\begin{aligned}\nabla_d \bar{L} &= \nabla_x^2 L(x^*, \mu^*) d + \nabla h(x^*) \bar{\mu} \\ \nabla_d^2 \bar{L} &= \nabla_x^2 L(x^*, \mu^*)\end{aligned}$$

Verifichiamo se $d = 0$, $\bar{\mu} = \mu^*$ è una coppia che soddisfa le C.N. di I e II ordine

$$\begin{aligned}\nabla_d \bar{L}(0, \mu^*) &= \nabla h(x^*) \mu^* \\ h(x^*) &= 0\end{aligned}$$

In generale **non** risulta $\nabla h(x^*) \mu^* = 0$. Si pensi a punti x^* "regolari" cioè tali per cui risultano lin. indipendenti le colonne di $\nabla h(x^*)$ e quindi $\nabla h(x^*) \mu = 0$ implica $\mu = 0 \neq \mu^*$.

Quindi, **non** valgono (in generale) le C.N. del I ordine, e quindi (x^*, μ^*) non è soluzione del problema approssimante

Approssimazione – 3° tentativo

- Sviluppo in serie di Taylor troncato al II ordine di $L(x, \mu^*)$ più $\nabla f(x^*)^\top (x - x^*)$
- Sviluppo in serie di Taylor troncato al I ordine di $h(x)$

$$\begin{array}{ll} \min_x & L(x^*, \mu^*) + \nabla f(x^*)^\top (x - x^*) + \frac{1}{2}(x - x^*)^\top \nabla_x^2 L(x^*, \mu^*)(x - x^*) \\ c.v. & h(x^*) + \nabla h(x^*)^\top (x - x^*) = 0 \end{array}$$

$$\begin{array}{ll} \min_d & L(x^*, \mu^*) + \nabla f(x^*)^\top d + \frac{1}{2} d^\top \nabla_x^2 L(x^*, \mu^*) d \\ c.v. & h(x^*) + \nabla h(x^*)^\top d = 0 \end{array}$$

Vogliamo essere sicuri che $(x^*, \mu^*) = (d^* = 0, \mu^*)$ sia soluzione anche del problema approssimante

Approssimazione – 3° tentativo

- Sviluppo in serie di Taylor troncato al II ordine di $L(x, \mu^*)$ più $\nabla f(x^*)^\top (x - x^*)$
- Sviluppo in serie di Taylor troncato al I ordine di $h(x)$

$$\begin{array}{ll} \min_x & L(x^*, \mu^*) + \nabla f(x^*)^\top (x - x^*) + \frac{1}{2}(x - x^*)^\top \nabla_x^2 L(x^*, \mu^*)(x - x^*) \\ c.v. & h(x^*) + \nabla h(x^*)^\top (x - x^*) = 0 \end{array}$$

$$\begin{array}{ll} \min_d & L(x^*, \mu^*) + \nabla f(x^*)^\top d + \frac{1}{2} d^\top \nabla_x^2 L(x^*, \mu^*) d \\ c.v. & h(x^*) + \nabla h(x^*)^\top d = 0 \end{array}$$

Vogliamo essere sicuri che $(x^*, \mu^*) = (d^* = 0, \mu^*)$ sia soluzione anche del problema approssimante

Approssimazione – 3° tentativo

$$
\begin{aligned}
\nabla_d \bar{L} &= \nabla_x^2 L(x^*, \mu^*) d + \nabla f(x^*) + \nabla h(x^*) \bar{\mu} \\
\nabla_d^2 \bar{L} &= \nabla_x^2 L(x^*, \mu^*)
\end{aligned}
$$

Verifichiamo se $d = 0$, $\bar{\mu} = \mu^*$ è una coppia che soddisfa le C.N. di I e II ordine

Approssimazione – 3° tentativo

$$\begin{aligned} \nabla_d \bar{L} &= \nabla_x^2 L(x^*, \mu^*) d + \nabla f(x^*) + \nabla h(x^*) \bar{\mu} \\ \nabla_d^2 \bar{L} &= \nabla_x^2 L(x^*, \mu^*) \end{aligned}$$

Verifichiamo se $d = 0$, $\bar{\mu} = \mu^*$ è una coppia che soddisfa le C.N. di I e II ordine

$$\begin{aligned} \nabla_d \bar{L} &= \nabla_x L(x^*, \mu^*) \\ h(x^*) + \nabla h(x^*)^\top d &= 0 \end{aligned}$$

Inoltre, dato che $\nabla_d^2 \bar{L} = \nabla_x^2 L(x^*, \mu^*)$ e per le ipotesi fatte sulla coppia (x^*, μ^*), quest'ultima verifica anche le C.S. del II ordine e quindi è **soluzione locale stretta** del problema approssimante

Approssimazione – 3° tentativo

$$
\begin{aligned}
\nabla_d \bar{L} &= \nabla_x^2 L(x^*, \mu^*)d + \nabla f(x^*) + \nabla h(x^*)\bar{\mu} \\
\nabla_d^2 \bar{L} &= \nabla_x^2 L(x^*, \mu^*)
\end{aligned}
$$

Verifichiamo se $d = 0$, $\bar{\mu} = \mu^*$ è una coppia che soddisfa le C.N. di I e II ordine

$$
\begin{aligned}
\nabla_d \bar{L}(0, \mu^*) &= 0 \qquad \text{C.N. I ordine OK} \\
h(x^*) &= 0
\end{aligned}
$$

Inoltre, dato che $\nabla_d^2 \bar{L} = \nabla_x^2 L(x^*, \mu^*)$ e per le ipotesi fatte sulla coppia (x^*, μ^*), quest'ultima verifica anche le C.S. del II ordine e quindi è **soluzione locale stretta** del problema approssimante

Approssimazione – 3° tentativo

$$\begin{aligned}
\nabla_d \bar{L} &= \nabla_x^2 L(x^*, \mu^*) d + \nabla f(x^*) + \nabla h(x^*) \bar{\mu} \\
\nabla_d^2 \bar{L} &= \nabla_x^2 L(x^*, \mu^*)
\end{aligned}$$

Verifichiamo se $d = 0$, $\bar{\mu} = \mu^*$ è una coppia che soddisfa le C.N. di I e II ordine

$$\begin{aligned}
\nabla_d \bar{L}(0, \mu^*) &= 0 \qquad \text{C.N. I ordine OK} \\
h(x^*) &= 0
\end{aligned}$$

Inoltre, dato che $\nabla_d^2 \bar{L} = \nabla_x^2 L(x^*, \mu^*)$ e per le ipotesi fatte sulla coppia (x^*, μ^*), quest'ultima verifica anche le C.S. del II ordine e quindi è **soluzione locale stretta** del problema approssimante

Introduzione

Consideriamo il problema

$$\begin{array}{ll} \min_x & f(x) \\ c.v. & h(x) = 0 \end{array}$$

dati (x_k, μ_k), le precedenti trasparenze suggeriscono di definire il problema "approssimante"

$$\begin{array}{ll} \min_d & L(x_k, \mu_k) + \nabla f(x_k)^\top d + \frac{1}{2} d^\top \nabla_x^2 L(x_k, \mu_k) d \\ c.v. & h(x_k) + \nabla h(x_k)^\top d = 0 \end{array}$$

Il metodo di Newton-Lagrange

Scriviamo (ancora) le condizioni di ottimo per il problema approssimante. $(\bar{d}, \bar{\mu})$ è soluzione quando:

- C.N. del I ordine

$$\nabla_x^2 L(x_k, \mu_k)\bar{d} + \nabla f(x_k) + \nabla h(x_k)\bar{\mu} = 0$$
$$\nabla h(x_k)^\top \bar{d} + h(x_k) = 0$$

- C.S. del II ordine

$$rg(\nabla h(x_k)) = p$$
$$d^\top \nabla_x^2 L(x_k, \mu_k) d > 0, \text{ per ogni } d \neq 0 \ t.c. \ \nabla h(x_k)^\top d = 0$$

Il metodo di Newton-Lagrange

Dalle C.N. del I ordine ricaviamo

$$\begin{bmatrix} \nabla^2_{xx} L(x_k, \mu_k) & \nabla h(x_k) \\ \nabla h(x_k)^\top & 0 \end{bmatrix} \begin{bmatrix} \bar{d} \\ \bar{\mu} \end{bmatrix} = - \begin{bmatrix} \nabla f(x_k) \\ h(x_k) \end{bmatrix}$$

La matrice dei coefficienti di questo sistema

$$M_k = \begin{bmatrix} \nabla^2_{xx} L(x_k, \mu_k) & \nabla h(x_k) \\ \nabla h(x_k)^\top & 0 \end{bmatrix}$$

è nota come matrice di KKT.

N.B. M_k è non singolare (e quindi invertibile) se valgono le C.S. del II ordine in k. Questo è vero se

- (x^*, μ^*) soddisfa le C.S. del II ordine
- (x_k, μ_k) è sufficientemente vicino a (x^*, μ^*)

Il metodo di Newton-Lagrange

Sia $(\bar{d}^k, \bar{\mu}^k)$ una soluzione del sistema, allora poniamo

$$\begin{aligned} x_{k+1} &= x_k + \bar{d}^k \\ \mu_{k+1} &= \bar{\mu}^k \end{aligned}$$

Metodo di soluzione (vincoli di uguaglianza)

Algoritmo `SQP`

Dati: (x_0, μ_0), `maxit`

for $k = 0, 1, \ldots,$`maxit`

Calcola $\bar{d}_k$ e $\bar{\mu}_k$ (inverti M_k)

Poni $x_{k+1} = x_k + \bar{d}_k$ e $\mu_{k+1} = \bar{\mu}_k$

if (x_{k+1}, μ_{k+1}) è KKT **then**

$x^* \leftarrow x_{k+1}$, $\mu^* \leftarrow \mu_{k+1}$ e STOP

endif

endfor

Return: miglior coppia trovata (x^*, μ^*)

Metodo di soluzione (vincoli di uguaglianza)

Algoritmo `SQP`

Dati: (x_0, μ_0), `maxit`

for $k = 0, 1, \ldots,$`maxit`

Calcola $\bar{d}_k$ e $\bar{\mu}_k$ (inverti M_k)

Poni $x_{k+1} = x_k + \bar{d}_k$ e $\mu_{k+1} = \bar{\mu}_k$

if (x_{k+1}, μ_{k+1}) è KKT **then**

$x^* \leftarrow x_{k+1}$, $\mu^* \leftarrow \mu_{k+1}$ e STOP

endif

endfor

Return: miglior coppia trovata (x^*, μ^*)

Metodo di soluzione (vincoli di uguaglianza)

Algoritmo `SQP`

Dati: (x_0, μ_0), `maxit`

for $k = 0, 1, \ldots,$`maxit`

Calcola $\bar{d}_k$ e $\bar{\mu}_k$ (inverti M_k)

Poni $x_{k+1} = x_k + \bar{d}_k$ e $\mu_{k+1} = \bar{\mu}_k$

if (x_{k+1}, μ_{k+1}) è KKT **then**

$x^* \leftarrow x_{k+1}$, $\mu^* \leftarrow \mu_{k+1}$ e STOP

endif

endfor

Return: miglior coppia trovata (x^*, μ^*)

Proprietà di convergenza di SQP

Proposizione

Sia (x^*, μ^*) soluzione del problema tale che siano soddisfatte

- *LICQ*[a]
- *SOSC*[b]

Allora, se (x_0, μ_0) è suff. vicino a (x^*, μ^*), l'algoritmo SQP *è ben definito e converge alla soluzione del problema*

[a] Linear Independence Constraint Qualification
[b] Second Order Sufficient Condition

Vincoli di disuguaglianza

Come si gestiscono in questo contesto vincoli di disuguaglianza $g(x) \leq 0$?

1) Aggiunta di variabili slack: $g_i(x) + s_i = 0$, $s_i \geq 0$ e gestione "esplicita" dei vincoli di bound sulla variabili (`SNOPT`, `KNITRO`)
2) Approccio S**i**QP - SQP with inequalities
3) Approccio S**e**QP - SQP with equalities (`SNOPT`)

Vincoli di disuguaglianza

$$\begin{array}{ll} \min_x & f(x) \\ c.v & h(x) = 0 \\ & g(x) \leq 0 \end{array}$$

Allora, data una stima corrente (x_k, μ_k, λ_k) della soluzione, definisci il problema

$$\begin{array}{ll} \min\limits_{d} & \nabla f(x_k)^\top d + \frac{1}{2} d^\top \nabla^2_{xx} L(x_k, \mu_k, \lambda_k) d \\ c.v. & \nabla h(x_k)^\top d + h(x_k) = 0 \\ & \nabla g(x_k)^\top d + g(x_k) \leq 0 \end{array}$$

N.B. ogni sottoproblema ha esattamente lo stesso numero di vincoli del problema originario

Vincoli di disuguaglianza

$$\begin{array}{ll} \min_x & f(x) \\ c.v & h(x) = 0 \\ & g(x) \leq 0 \end{array}$$

Allora, data una stima corrente (x_k, μ_k, λ_k) della soluzione, definisci il problema

$$\begin{array}{ll} \min\limits_{d} & \nabla f(x_k)^\top d + \frac{1}{2} d^\top \nabla^2_{xx} L(x_k, \mu_k, \lambda_k) d \\ c.v. & \nabla h(x_k)^\top d + h(x_k) = 0 \\ & \nabla g(x_k)^\top d + g(x_k) \leq 0 \end{array}$$

N.B. ogni sottoproblema ha esattamente lo stesso numero di vincoli del problema originario

Vincoli di disuguaglianza

$$\begin{array}{ll} \min_x & f(x) \\ c.v & h(x) = 0 \\ & g(x) \leq 0 \end{array}$$

Immaginiamo di poter conoscere quali vincoli di disuguaglianza sono attivi in una soluzione del problema x^*.
Indichiamo I^* l'insieme degli indici dei vincoli attivi in x^*

$$I^* = \{i : g_i(x^*) = 0\}$$

Quindi, $g_i(x^*) < 0$, per ogni $i \notin I^*$

Vincoli di disuguaglianza

$$\begin{array}{ll} \min_x & f(x) \\ c.v & h(x) = 0 \\ & g(x) \leq 0 \end{array}$$

Immaginiamo di poter conoscere quali vincoli di disuguaglianza sono attivi in una soluzione del problema x^*.
Indichiamo I^* l'insieme degli indici dei vincoli attivi in x^*

$$I^* = \{i : g_i(x^*) = 0\}$$

Quindi, $g_i(x^*) < 0$, per ogni $i \notin I^*$

Vincoli di disuguaglianza

Sotto questa ipotesi, anche il problema

$$\begin{array}{ll} \min_x & f(x) \\ c.v & h(x) = 0 \\ & g_i(x) = 0, \quad i \in I^* \end{array}$$

ammette x^* come soluzione locale.

N.B. questo problema ha (nuovamente) tutti e soli vincoli di uguaglianza. Pertanto, posso usare il metodo di soluzione che abbiamo visto per problemi con vincoli di uguaglianza

Vincoli di disuguaglianza

Sotto questa ipotesi, anche il problema

$$\begin{array}{ll} \min_x & f(x) \\ c.v & h(x) = 0 \\ & g_i(x) = 0, \quad i \in I^* \end{array}$$

ammette x^* come soluzione locale.
N.B. questo problema ha (nuovamente) tutti e soli vincoli di uguaglianza. Pertanto, posso usare il metodo di soluzione che abbiamo visto per problemi con vincoli di uguaglianza

Vincoli di disuguaglianza

Cosa manca? Normalmente non sappiamo come è fatto I^*. Quindi:

- si dà una stima I_k di I^*
- si risolve il sotto problema
- si aggiorna la stima di I^* definendo I_{k+1}

Per questo motivo questa classe di metodi è anche nota come "active-set" SQP

Vincoli di disuguaglianza

Cosa manca? Normalmente non sappiamo come è fatto I^*. Quindi:

- si dà una stima I_k di I^*
- si risolve il sotto problema
- si aggiorna la stima di I^* definendo I_{k+1}

Per questo motivo questa classe di metodi è anche nota come "active-set" SQP

Vincoli di disuguaglianza

Cosa manca? Normalmente non sappiamo come è fatto I^*. Quindi:

- si dà una stima I_k di I^*
- si risolve il sotto problema
- si aggiorna la stima di I^* definendo I_{k+1}

Per questo motivo questa classe di metodi è anche nota come "active-set" SQP

Problema approssimante

Data una stima dei vincoli attivi I_k ed una stima della soluzione (x_k, μ_k, λ_k), il problema "approssimante" è

$$\begin{array}{ll} \min\limits_{d} & \nabla f(x_k)^\top d + \dfrac{1}{2} d^\top \nabla^2_{xx} L(x_k, \mu_k, \lambda_k) d \\ c.v. & \nabla h(x_k)^\top d + h(x_k) = 0 \\ & \nabla g_{I_k}(x_k)^\top d + g_{I_k}(x_k) = 0 \end{array}$$

Soluzione del problema approssimante

In base a ragionamenti del tutto analoghi a quelli appena visti, bisogna risolvere il sistema lineare:

$$\begin{bmatrix} \nabla^2_{xx} L_k & \nabla h(x_k) & \nabla g_{I_k}(x_k) \\ \nabla h(x_k)^\top & 0 & 0 \\ \nabla g_{I_k}(x_k)^\top & 0 & 0 \end{bmatrix} \begin{bmatrix} \bar{d} \\ \bar{\mu} \\ \bar{\lambda} \end{bmatrix} = - \begin{bmatrix} \nabla f(x_k) \\ h(x_k) \\ g_{I_k}(x_k) \end{bmatrix}$$

Metodo di soluzione (vincoli di uguaglianza)

Algoritmo `SQP`

Dati: (x_0, μ_0, λ_0), `maxit`, $\gamma > 0$

for $k = 0, 1, \dots,$`maxit`

$I_k = \{i : g_i(x_k) \geq -\gamma(\lambda_k)_i\}$

Calcola $\bar{d}_k$, $\bar{\mu}_k$, $\bar{\lambda}_k$

Poni $x_{k+1} = x_k + \bar{d}_k$, $\mu_{k+1} = \bar{\mu}_k$,
$(\lambda_{k+1})_{I_k} = \bar{\lambda}_k$, $(\lambda_{k+1})i = 0$, $i \notin I_k$

if $(x_{k+1}, \mu_{k+1}, \lambda_{k+1})$ è KKT **then**

$x^* \leftarrow x_{k+1}$, $\mu^* \leftarrow \mu_{k+1}$, $\lambda^* \leftarrow \lambda_{k+1}$ e STOP

endif

endfor

Return: miglior punto trovato (x^*, μ^*, λ^*)

Metodo di soluzione (vincoli di uguaglianza)

Algoritmo `SQP`

Dati: (x_0, μ_0, λ_0), `maxit`, $\gamma > 0$

for $k = 0, 1, \dots,$`maxit`

$I_k = \{i : g_i(x_k) \geq -\gamma(\lambda_k)_i\}$

Calcola $\bar{d}_k$, $\bar{\mu}_k$, $\bar{\lambda}_k$

Poni $x_{k+1} = x_k + \bar{d}_k$, $\mu_{k+1} = \bar{\mu}_k$,
$(\lambda_{k+1})_{I_k} = \bar{\lambda}_k$, $(\lambda_{k+1})i = 0$, $i \notin I_k$

if $(x_{k+1}, \mu_{k+1}, \lambda_{k+1})$ è KKT **then**

$x^* \leftarrow x_{k+1}$, $\mu^* \leftarrow \mu_{k+1}$, $\lambda^* \leftarrow \lambda_{k+1}$ e STOP

endif

endfor

Return: miglior punto trovato (x^*, μ^*, λ^*)

Metodo di soluzione (vincoli di uguaglianza)

Algoritmo `SQP`

Dati: (x_0, μ_0, λ_0), `maxit`, $\gamma > 0$

for $k = 0, 1, \dots,$`maxit`

$I_k = \{i : g_i(x_k) \geq -\gamma(\lambda_k)_i\}$

Calcola $\bar{d}_k$, $\bar{\mu}_k$, $\bar{\lambda}_k$

Poni $x_{k+1} = x_k + \bar{d}_k$, $\mu_{k+1} = \bar{\mu}_k$,
$(\lambda_{k+1})_{I_k} = \bar{\lambda}_k$, $(\lambda_{k+1})i = 0$, $i \notin I_k$

if $(x_{k+1}, \mu_{k+1}, \lambda_{k+1})$ è KKT **then**

$x^* \leftarrow x_{k+1}$, $\mu^* \leftarrow \mu_{k+1}$, $\lambda^* \leftarrow \lambda_{k+1}$ e STOP

endif

endfor

Return: miglior punto trovato (x^*, μ^*, λ^*)