# Ottimizzazione dei Sistemi Complessi

G. Liuzzi[1]

Giovedì 27 Aprile 2017

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

# Introduzione

Consideriamo il problema

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

dati $(x_k, \mu_k)$, si definisce il problema “approssimante”

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

# 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

- $(x^*, \mu^*)$ soddisfa le C.S. del II ordine
- $(x_k, \mu_k)$ è sufficientemente vicino a $(x^*, \mu^*)$

# 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**: $(x_0, \mu_0)$, `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** $(x_{k+1}, \mu_{k+1})$ è KKT **then**

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

**endif**

**endfor**

**Return**: miglior coppia trovata $(x^*, \mu^*)$

# Proprietà di convergenza di SQP

## Proposizione

*Sia* $(x^*, \mu^*)$ *soluzione del problema tale che siano soddisfatte*

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

*Allora, se* $(x_0, \mu_0)$ *è suff. vicino a* $(x^*, \mu^*)$*, 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 $(x_k, \mu_k, \lambda_k)$ della soluzione, definisci il problema

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

**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

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

# Problema approssimante

Data una stima dei vincoli attivi $I_k$ ed una stima della soluzione $(x_k, \mu_k, \lambda_k)$, il problema "approssimante" è

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

# 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**: $(x_0, \mu_0, \lambda_0)$, $\texttt{maxit}$, $\gamma > 0$

**for** $k = 0, 1, \ldots, \texttt{maxit}$

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

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 $(x^*, \mu^*, \lambda^*)$