# Ottimizzazione dei Sistemi Complessi

G. Liuzzi[1]

Giovedì 26 Aprile 2018

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

# Introduzione

Consideriamo il problema (con soli vincoli di **disuguaglianza**)

$$\begin{array}{ll} \min\limits_{x} & f(x) \\ c.v. & g(x) \geq 0 \end{array}$$

- $\mathcal{F}$ regione ammissibile
- $\mathring{\mathcal{F}} = Int(\mathcal{F}) = \{x \in \mathbb{R}^n : g_i(x) > 0\}$
- Assumiamo $\mathring{\mathcal{F}} \neq \emptyset$

# Funzione di Barriera

La più diffusa funzione di "barriera" per il problema considerato è:

$$
\begin{aligned}
P(x;\mu) &= f(x) - \mu \sum_{i=1}^{m} \log g_i(x), \ \text{dove} \\
b(x) &= -\sum_{i=1}^{m} \log g_i(x) \ \text{(termine di barriera)} \\
\nabla_x P(x;\mu) &= \nabla f(x) - \sum_{i=1}^{m} \frac{\mu}{g_i(x)} \nabla g_i(x)
\end{aligned}
$$

**N.B.**

- $b(x)$ (e quindi $P(x;\mu)$) è definita per ogni $x \in \mathring{\mathcal{F}}$
- Supponiamo $b(x)$ (e quindi $P(x;\mu)$) $= +\infty$ per ogni $x \notin \mathring{\mathcal{F}}$

# Condizioni di KKT – p.ti staz. di $P(x;\mu)$

$\boxed{\text{P.ti di KKT}}$

Sia $(x,\lambda)$ tale che

$$
\begin{aligned}
\nabla f(x) - \lambda^\top \nabla g(x) &= 0 \\
g(x) &\geq 0 \\
\lambda &\geq 0 \\
\lambda^\top g(x) &= 0
\end{aligned}
$$

$\boxed{\text{P.ti staz. di } P(x;\mu)}$

Sia $x(\mu)$ tale che

$$\nabla f(x(\mu)) - \sum_{i=1}^{m} \lambda_i(\mu) \nabla g_i(x(\mu)) = 0$$

$$\lambda_i(\mu) = \frac{\mu}{g_i(x(\mu))}$$

- $(x(\mu), \lambda(\mu))$ definiscono il **central-path**
- soddisfano tutte le condizioni di KKT **tranne** $\lambda^\top g(x) = 0$
- però risulta:

$$\lambda_i(\mu) g_i(x(\mu)) = \mu$$

# Condizioni di KKT – p.ti staz. di $P(x;\mu)$

$\boxed{\text{P.ti di KKT}}$

Sia $(x, \lambda)$ tale che

$$\begin{array}{rcl} \nabla f(x) - \lambda^\top \nabla g(x) & = & 0 \\ g(x) & \geq & 0 \\ \lambda & \geq & 0 \\ \lambda^\top g(x) & = & 0 \end{array}$$

$\boxed{\text{P.ti staz. di } P(x;\mu)}$

Sia $x(\mu)$ tale che

$$\nabla f(x(\mu)) - \sum_{i=1}^{m} \lambda_i(\mu) \nabla g_i(x(\mu)) = 0$$

$$\lambda_i(\mu) = \frac{\mu}{g_i(x(\mu))}$$

- $(x(\mu), \lambda(\mu))$ definiscono il **central-path**
- soddisfano tutte le condizioni di KKT **tranne** $\lambda^\top g(x) = 0$
- però risulta:

$$\lambda_i(\mu) g_i(x(\mu)) = \mu$$

# Metodo di soluzione

**Algoritmo** `LOG-BARRIER`

**Dati**: $\mu_\circ > \mu_{tol} > 0$, $\tau_0 > 0$, $\mathtt{maxit}$, $x_0$ t.c. $g(x_0) > 0$

**for** $k = 0, 1, \dots, \mathtt{maxit}$

Calcola $x_k = x(\mu_k)$ t.c. $\|\nabla P(x_k; \mu_k)\| \leq \tau_k$

**if** $\mu_k < \mu_{tol}$ **then**

$x^* \leftarrow x_k$, $\lambda^* \leftarrow \mu_k / g(x_k)$, STOP

**endif**

**endfor**

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

# Metodo di soluzione

**Algoritmo** `LOG-BARRIER`

**Dati**: $\mu_\circ > \mu_{tol} > 0$, $\tau_0 > 0$, `maxit`, $x_0$ t.c. $g(x_0) > 0$

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

Calcola $x_k = x(\mu_k)$ t.c. $\|\nabla P(x_k; \mu_k)\| \leq \tau_k$

**if** $\mu_k < \mu_{tol}$ **then**

$x^* \leftarrow x_k$, $\lambda^* \leftarrow \mu_k / g(x_k)$, STOP

**endif**

Scegli $\mu_{k+1} \in (0, \mu_k)$

**endfor**

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

# Metodo di soluzione

**Algoritmo** `LOG-BARRIER`

**Dati**: $\mu_\circ > \mu_{tol} > 0$, $\tau_0 > 0$, $\mathtt{maxit}$, $x_0$ t.c. $g(x_0) > 0$

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

Calcola $x_k = x(\mu_k)$ t.c. $\|\nabla P(x_k; \mu_k)\| \leq \tau_k$

**if** $\mu_k < \mu_{tol}$ **then**

$x^* \leftarrow x_k$, $\lambda^* \leftarrow \mu_k / g(x_k)$, STOP

**endif**

Scegli $\mu_{k+1} \in (0, \mu_k)$

**endfor**

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

# Metodo di soluzione

**Algoritmo** `LOG-BARRIER`

**Dati**: $\mu_\circ > \mu_{tol} > 0$, $\tau_0 > 0$, $\texttt{maxit}$, $x_0$ t.c. $g(x_0) > 0$

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

Calcola $x_k = x(\mu_k)$ t.c. $\|\nabla P(x_k; \mu_k)\| \leq \tau_k$

**if** $\mu_k < \mu_{tol}$ **then**

$x^* \leftarrow x_k$, $\lambda^* \leftarrow \mu_k / g(x_k)$, STOP

**endif**

Scegli $\mu_{k+1} \in (0, \mu_k)$

**endfor**

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

# Teoricamente ...

## Proposizione

*$\mathring{\mathcal{F}} \neq \emptyset$, $(x^*, \lambda^*)$ soluzione locale del problema vincolato che soddisfa LICQ[a], stretta complementarità e SOSC[b]. Allora:*

- $\nabla^2_{xx} P(x;\mu)$ *è definita positiva per $\mu$ suff. piccolo*
- *esiste, unica e cont. differenziabile la funzione $x(\mu)$ (minimo locale di $P$ in un intorno di $x^*$ e per $\mu$ suff. piccoli)*
- $\lim_{\mu \downarrow 0} x(\mu) = x^*$
- $\lim_{\mu \downarrow 0} \lambda(\mu) = \lambda^*$

[a] Linear Independence Constraint Qualification

[b] Second Order Sufficient Condition

Cioè `LOG-BARRIER`, sotto opportune ipotesi, converge alla coppia di KKT $(x^*, \lambda^*)$

# Teoricamente ...

## Proposizione

*$\overset{\circ}{\mathcal{F}} \neq \emptyset$, $(x^*, \lambda^*)$ soluzione locale del problema vincolato che soddisfa LICQ[a], stretta complementarità e SOSC[b]. Allora:*

- $\nabla^2_{xx} P(x;\mu)$ *è definita positiva per $\mu$ suff. piccolo*
- *esiste, unica e cont. differenziabile la funzione $x(\mu)$ (minimo locale di $P$ in un intorno di $x^*$ e per $\mu$ suff. piccoli)*
- $\lim_{\mu \downarrow 0} x(\mu) = x^*$
- $\lim_{\mu \downarrow 0} \lambda(\mu) = \lambda^*$

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

Cioè `LOG-BARRIER`, sotto opportune ipotesi, converge alla coppia di KKT $(x^*, \lambda^*)$

# Vincoli di uguaglianza

Consideriamo il problema

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

La funzione di “barriera” è in questo caso:

$$P(x;\mu,\epsilon) = f(x) - \mu \sum_{i=1}^{m} \log g_i(x) + \frac{1}{\epsilon}\|h(x)\|^2$$

oppure

$$\begin{aligned} P(x;\mu) &= f(x) - \mu \sum_{i=1}^{m} \log g_i(x) + \frac{1}{2\mu}\|h(x)\|^2 \\ \nabla_x P(x;\mu) &= \nabla f(x) - \sum_{i=1}^{m} \frac{\mu}{g_i(x)} \nabla g_i(x) + \frac{1}{\mu} \sum_{i=1}^{p} h_i(x) \nabla h_i(x) \end{aligned}$$

# Vincoli di uguaglianza

Consideriamo il problema

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

La funzione di "barriera" è in questo caso:

$$P(x;\mu,\epsilon) = f(x) - \mu \sum_{i=1}^{m} \log g_i(x) + \frac{1}{\epsilon}\|h(x)\|^2$$

oppure

$$\begin{aligned} P(x;\mu) &= f(x) - \mu \sum_{i=1}^{m} \log g_i(x) + \frac{1}{2\mu}\|h(x)\|^2 \\ \nabla_x P(x;\mu) &= \nabla f(x) - \sum_{i=1}^{m} \frac{\mu}{g_i(x)} \nabla g_i(x) + \frac{1}{\mu}\sum_{i=1}^{p} h_i(x)\nabla h_i(x) \end{aligned}$$

# Punto iniziale

Siccome,

$$P(x;\mu) < +\infty \text{ se e solo se } g(x) > 0,$$

allora ...

per poter utilizzare la funzione $P(x;\mu)$ in un algoritmo di soluzione per il problema vincolato è **necessario** conoscere un punto $x_0$ tale che $g(x_0) > 0$ (cfr. `LOG-BARRIER`).

# Punto iniziale

Siccome,

$$P(x;\mu) < +\infty \text{ se e solo se } g(x) > 0,$$

allora ...

per poter utilizzare la funzione $P(x;\mu)$ in un algoritmo di soluzione per il problema vincolato è **necessario** conoscere un punto $x_0$ tale che $g(x_0) > 0$ (cfr. `LOG-BARRIER`).

# Variabili slack

$$\begin{array}{ll} \min\limits_{x} & f(x) \\ c.v. & g(x) \geq 0 \end{array} \quad \rightsquigarrow \quad \begin{array}{ll} \min\limits_{x,s} & f(x) \\ c.v. & g_i(x) - s_i = 0 \\ & \boxed{s_i \geq 0} \end{array}$$

$$P(x, s; \mu) = f(x) - \mu \sum_{i=1}^{m} \log s_i + \frac{1}{2\mu} \sum_{i=1}^{m} (g_i(x) - s_i)^2$$

$\boxed{\text{Metodi interni} \rightarrow \text{inammissibili}}$

**N.B.** $P(x; \mu)$ è definita anche quando $g(x) \not> 0$

# Variabili slack

$$\begin{array}{ll} \min\limits_{x} & f(x) \\ c.v. & g(x) \geq 0 \end{array} \quad \rightsquigarrow \quad \begin{array}{ll} \min\limits_{x,s} & f(x) \\ c.v. & g_i(x) - s_i = 0 \\ & \boxed{s_i \geq 0} \end{array}$$

$$P(x, s; \mu) = f(x) - \mu \sum_{i=1}^{m} \log s_i + \frac{1}{2\mu} \sum_{i=1}^{m} (g_i(x) - s_i)^2$$

$\boxed{\text{Metodi interni – inammissibili}}$

**N.B.** $P(x; \mu)$ è definita anche quando $g(x) \not\geq 0$

# Variabili slack

$$\begin{array}{ll} \min\limits_{x} & f(x) \\ c.v. & g(x) \geq 0 \end{array} \quad \rightsquigarrow \quad \begin{array}{ll} \min\limits_{x,s} & f(x) \\ c.v. & g_i(x) - s_i = 0 \\ & \boxed{s_i \geq 0} \end{array}$$

$$P(x,s;\mu) = f(x) - \mu \sum_{i=1}^{m} \log s_i + \frac{1}{2\mu} \sum_{i=1}^{m} (g_i(x) - s_i)^2$$

Metodi interni – inammissibili

**N.B.** $P(x;\mu)$ è definita anche quando $g(x) \not\geq 0$

# Variabili slack

$$\begin{array}{ll} \min\limits_{x} & f(x) \\ c.v. & g(x) \geq 0 \end{array} \quad \rightsquigarrow \quad \begin{array}{ll} \min\limits_{x,s} & f(x) \\ c.v. & g_i(x) - s_i = 0 \\ & \boxed{s_i \geq 0} \end{array}$$

$$P(x, s; \mu) = f(x) - \mu \sum_{i=1}^{m} \log s_i + \frac{1}{2\mu} \sum_{i=1}^{m} (g_i(x) - s_i)^2$$

$\boxed{\text{Metodi interni – inammissibili}}$

**N.B.** $P(x; \mu)$ è definita anche quando $g(x) \not\geq 0$

## Relazione con KKT

$$\begin{array}{ll} \min\limits_{x,s} & f(x) \\ c.v. & g_i(x) - s_i = 0 \\ & s_i \geq 0 \end{array}$$

Funzione Lagrangiana: $L(x, \rho, \lambda) = f(x) + \rho^\top (g - s) - \lambda^\top s$

Cond. di KKT:

$$\begin{aligned} \nabla_x L &= \nabla f(x) + \sum_{i=1}^{m} \rho_i \nabla g_i(x) = 0 \\ \nabla_s L &= -\rho - \lambda = 0 \Rightarrow \rho = -\lambda \\ g - s &= 0 \\ \lambda &\geq 0,\ s \geq 0,\ \lambda^\top s = 0 \end{aligned}$$

# Relazione con KKT

$$\begin{array}{ll} \min\limits_{x,s} & f(x) \\ c.v. & g_i(x) - s_i = 0 \\ & s_i \geq 0 \end{array}$$

Funzione Lagrangiana: $L(x, \rho, \lambda) = f(x) + \rho^\top(g - s) - \lambda^\top s$

Cond. di KKT:

$$\begin{array}{rcl} \nabla_x L & = & \nabla f(x) + \sum\limits_{i=1}^{m} \rho_i \nabla g_i(x) = 0 \\ \nabla_s L & = & -\rho - \lambda = 0 \Rightarrow \rho = -\lambda \\ g - s & = & 0 \\ \lambda & \geq & 0,\ s \geq 0,\ \lambda^\top s = 0 \end{array}$$

# Relazione con KKT

$$\begin{aligned}\min_{x,s} \quad & f(x) \\ c.v. \quad & g_i(x) - s_i = 0 \\ & s_i \geq 0\end{aligned}$$

Funzione Lagrangiana: $L(x,\rho,\lambda) = f(x) + \rho^\top (g - s) - \lambda^\top s$

Cond. di KKT:

$$\begin{aligned}\nabla_x L &= \nabla f(x) - \sum_{i=1}^{m} \lambda_i \nabla g_i(x) = 0 \\ \nabla_s L &= -\rho - \lambda = 0 \Rightarrow \rho = -\lambda \\ g - s &= 0 \\ \lambda &\geq 0,\ s \geq 0,\ \lambda^\top s = 0\end{aligned}$$

# Relazione con KKT

$$\min_{x,s} \quad P(x, s; \mu)$$

$$
\begin{aligned}
P(x, s; \mu) &= f(x) - \mu \sum_{i=1}^{m} \log s_i + \frac{1}{2\mu} \sum_{i=1}^{m} (g_i(x) - s_i)^2 \\
\nabla_x P &= \nabla f(x) + \sum_{i=1}^{m} \frac{g_i(x) - s_i}{\mu} \nabla g_i(x) = 0 \\
\nabla_{s_i} P &= -\frac{\mu}{s_i} - \frac{g_i(x) - s_i}{\mu} = 0 \Rightarrow \frac{g_i(x) - s_i}{\mu} = -\frac{\mu}{s_i} = -\lambda_i \\
\lambda_i s_i &= \mu, \quad g_i(x) - s_i = -\frac{\mu^2}{s_i} = -\lambda_i \mu
\end{aligned}
$$

# Relazione con KKT

$$\min_{x,s} \quad P(x,s;\mu)$$

$$\begin{aligned}
P(x,s;\mu) &= f(x) - \mu \sum_{i=1}^{m} \log s_i + \frac{1}{2\mu} \sum_{i=1}^{m} (g_i(x) - s_i)^2 \\
\nabla_x P &= \nabla f(x) + \sum_{i=1}^{m} \frac{g_i(x) - s_i}{\mu} \nabla g_i(x) = 0 \\
\nabla_{s_i} P &= -\frac{\mu}{s_i} - \frac{g_i(x) - s_i}{\mu} = 0 \Rightarrow \frac{g_i(x) - s_i}{\mu} = -\frac{\mu}{s_i} = -\lambda_i \\
\lambda_i s_i &= \mu, \quad g_i(x) - s_i = -\frac{\mu^2}{s_i} = -\lambda_i \mu
\end{aligned}$$

# Relazione con KKT

$$\min_{x,s} \quad P(x,s;\mu)$$

$$
\begin{aligned}
P(x,s;\mu) &= f(x) - \mu \sum_{i=1}^{m} \log s_i + \frac{1}{2\mu} \sum_{i=1}^{m} (g_i(x) - s_i)^2 \\
\nabla_x P &= \nabla f(x) - \sum_{i=1}^{m} \lambda_i \nabla g_i(x) = 0 \\
\nabla_{s_i} P &= -\frac{\mu}{s_i} - \frac{g_i(x) - s_i}{\mu} = 0 \Rightarrow \frac{g_i(x) - s_i}{\mu} = -\frac{\mu}{s_i} = -\lambda_i \\
\lambda_i s_i &= \mu, \quad g_i(x) - s_i = -\frac{\mu^2}{s_i} = -\lambda_i \mu
\end{aligned}
$$