Ottimizzazione dei Sistemi Complessi

G. Liuzzi[1]

Giovedì 17 Marzo 2016

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

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

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

Sia (x,λ) 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:

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

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

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

endfor

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

Teoricamente ...

Proposizione

$\mathring{\mathcal{F}} \neq \emptyset$, (x^*, λ^*) 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 μ suff. piccolo*
- *esiste, unica e cont. differenziabile la funzione $x(\mu)$ (minimo locale di P in un intorno di x^* e per μ 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^*, λ^*)

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`).

Variabili slack

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

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

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

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

Relazione con KKT

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

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

Cond. di KKT:

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

Relazione con KKT

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

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

Introduzione

Consideriamo il problema

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

Le condizioni di KKT per il problema danno luogo al sistema nonlineare

$$F(x,\mu) = \left[\begin{array}{c} \nabla f(x) + \nabla h(x)\mu \\ h(x) \end{array} \right] = 0$$

Idea: Risolvere il sistema $F(x,\mu)$ usando il **metodo di Newton**

Metodo di Newton-Lagrange

Se vogliamo usare il metodo di Newton, la prima cosa da fare è calcolare lo Jacobiano di F

$$\begin{bmatrix} \nabla^2_{xx} L(x,\mu) & \nabla h(x) \\ \nabla h(x)^\top & 0 \end{bmatrix}$$

Il metodo di Newton è definito dalla iterazione

$$\begin{pmatrix} x_{k+1} \\ \mu_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ \mu_k \end{pmatrix} + \begin{pmatrix} \delta_k^x \\ \delta_k^\mu \end{pmatrix}$$

dove $(\delta_k^x, \delta_k^\mu)$ tali che

$$\begin{bmatrix} \nabla^2_{xx} L(x_k,\mu_k) & \nabla h(x_k) \\ \nabla h(x_k)^\top & 0 \end{bmatrix} \begin{pmatrix} \delta_k^x \\ \delta_k^\mu \end{pmatrix} = -F(x_k,\mu_k)$$

Metodo di Newton-Lagrange

Il passo k è ben definito quando la matrice

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

è **non singolare**. Questo accade sotto la seguente ipotesi:

Assunzione

- *$rg(\nabla h(x_k)) = p$, i.e. i gradienti dei vincoli in x_k sono lin. indipendenti*
- *$d^\top \nabla^2_{xx} L(x_k, \mu_k) d > 0$ per ogni $d \neq 0$ e tale che $\nabla h(x_k)^\top d = 0$*

Metodo SQP

È possibile vedere il metodo di Newton-Lagrange in maniera alternativa. A tal fine, definiamo il problema

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

Se vale l'assunzione di prima, questo problema ha una unica soluzione $(\bar{d}, \bar{\mu})$ che (ovviamente) soddisfa

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

Metodo SQP

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

lo possiamo scrivere come

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

e, sottraendo $\nabla h(x_k)\mu_k$ dalla prima eq.

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

Metodo SQP

Ponendo $\bar{d}_\mu = \bar{\mu} - \mu_k$ otteniamo proprio l'iterazione del metodo di Newton-Lagrange

$$\begin{pmatrix} x_{k+1} \\ \mu_{k+1} \end{pmatrix} = \begin{pmatrix} \bar{d} \\ \bar{d}_\mu \end{pmatrix} + \begin{pmatrix} x_k \\ \mu_k \end{pmatrix}$$

dove risulta $\mu_{k+1} = \bar{\mu}$

Metodo di soluzione

Algoritmo `SQP`

Dati: (x_0, μ_0), `maxit`

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

Calcola $\bar{d}_k$ e $\bar{\mu}_k$

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^*)

Vincoli di disuguaglianza

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

Allora, definisci il problema

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

Proprietà di convergenza di SQP

Proposizione

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

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

Allora, se (x_0, μ_0, λ_0) *è suff. vicino a* (x^*, μ^*, λ^*), *l'algoritmo* SQP *è ben definito e converge alla soluzione del problema*

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