Ottimizzazione dei Sistemi Complessi

G. Liuzzi[1]

Martedì 15 Marzo 2016

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

Problemi con soli vincoli di disuguaglianza – 1

Consideriamo il problema

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

Abbiamo due alternative:

1) $g_i(x) \leq 0 \quad \rightsquigarrow \quad g_i(x) + y_i^2 = 0,\ i = 1, \dots, m$

$$\begin{aligned} L_a(x, y, \lambda; \epsilon) &= f(x) + \sum_{i=1}^m \lambda_i (g_i(x) + y_i^2) + \frac{1}{\epsilon} \sum_{i=1}^m (g_i(x) + y_i^2)^2 \\ &= f(x) + \lambda^\top g(x) + \frac{1}{\epsilon} \|g(x)\|^2 + \frac{1}{\epsilon} \sum_{i=1}^m \left(\epsilon \lambda_i y_i^2 + 2 g_i(x) y_i^2 + y_i^4 \right) \end{aligned}$$

2) $g_i(x) \leq 0 \quad \rightsquigarrow \quad g_i(x) + s_i = 0,\ s_i \geq 0,\ i = 1, \dots, m$

$$L_a(x, s, \lambda; \epsilon) = f(x) + \sum_{i=1}^m \lambda_i (g_i(x) + s_i) + \frac{1}{\epsilon} \sum_{i=1}^m (g_i(x) + s_i)^2$$

Problemi con soli vincoli di disuguaglianza – 2

Se consideriamo l'**opzione 1**, quindi

$$\begin{array}{ll} \min & f(x) \\ c.v. & g(x) \leq 0 \end{array} \quad \text{ovvero} \quad \begin{array}{ll} \min & f(x) \\ c.v. & g_i(x) + y_i^2 = 0, \ i = 1, \dots, m \end{array}$$

abbiamo

$$L_a(x, y, \lambda; \epsilon) = f(x) + \lambda^\top g(x) + \frac{1}{\epsilon}\|g(x)\|^2 + \frac{1}{\epsilon}\sum_{i=1}^{m}\left(\epsilon\lambda_i y_i^2 + 2g_i(x)y_i^2 + y_i^4\right)$$

fissati ϵ e λ, bisogna "risolvere"

$$\min_{x,y} L_a(x, y, \lambda; \epsilon)$$

Problemi con soli vincoli di disuguaglianza – 3

Se consideriamo l'**opzione 2**, quindi

$$\begin{array}{ll} \min & f(x) \\ c.v. & g(x) \le 0 \end{array} \qquad \text{ovvero} \qquad \begin{array}{ll} \min & f(x) \\ c.v. & g_i(x) + s_i = 0, \;\; i = 1, \dots, m \\ & s_i \ge 0, \;\; i = 1, \dots, m \end{array}$$

abbiamo

$$L_a(x, s, \lambda; \epsilon) = f(x) + \sum_{i=1}^{m} \lambda_i (g_i(x) + s_i) + \frac{1}{\epsilon} \sum_{i=1}^{m} (g_i(x) + s_i)^2$$

fissati ϵ e λ, bisogna "risolvere"

$$\begin{array}{ll} \min\limits_{x,s} & L_a(x, s, \lambda; \epsilon) \\ c.v. & s \ge 0 \end{array}$$

Gradiente della Lagrangiana

$$
\begin{aligned}
\nabla_x L_a &= \nabla f(x) + \nabla g(x)\lambda + \frac{2}{\epsilon}\sum_{i=1}^{m}(g_i(x) + y_i^2)\nabla g_i(x) \\
\nabla_\lambda L_a &= g(x) + y^2 \\
\nabla_{y_i} L_a &= \frac{4}{\epsilon} y_i \left(\epsilon\frac{\lambda_i}{2} + g_i(x) + y_i^2\right)
\end{aligned}
$$

Quindi, ponendo $\nabla_{y_i} L_a = 0$ otteniamo

$$
y_i\left(\epsilon\frac{\lambda_i}{2} + g_i(x) + y_i^2\right) = 0
$$

ovvero

$$
y_i^2 = \begin{cases} 0 & casoA \\ \max\left\{0, -\epsilon\frac{\lambda_i}{2} - g_i(x)\right\} & casoB \end{cases}
$$

Vogliamo minimizzare

quindi, i termini

$$\left(\epsilon\lambda_i y_i^2 + 2g_i(x)y_i^2 + y_i^4\right)$$

valgono:

$$\begin{cases} 0 & casoA \\ 2\max\{0, -\epsilon\frac{\lambda_i}{2} - g_i\}(\epsilon\frac{\lambda_i}{2} + g_i) + \max\{0, -\epsilon\frac{\lambda_i}{2} - g_i\}^2 & casoB \end{cases}$$

Nel caso B, in particolare quando $-\epsilon\frac{\lambda_i}{2} - g_i > 0$, si ottiene

$$-\left(\epsilon\frac{\lambda_i}{2} + g_i(x)\right)^2 < 0$$

Quindi si può concludere che il minimo rispetto alle y_i lo si ottiene per

$$y_i^2 = \max\left\{0, -\epsilon\frac{\lambda_i}{2} - g_i\right\}.$$

Espressione per vincoli di disuguaglianza

$$
\begin{aligned}
L_a(x,\lambda;\epsilon) &= f(x) + \lambda^\top \left(g(x) + \max\left\{0, -\epsilon\frac{\lambda}{2} - g(x)\right\}\right) \\
&+ \frac{1}{\epsilon}\left\| g(x) + \max\left\{0, -\epsilon\frac{\lambda}{2} - g(x)\right\}\right\|^2
\end{aligned}
$$

Ovvero

$$
L_a(x,\lambda;\epsilon) = f(x) + \lambda^\top \max\left\{g(x), -\epsilon\frac{\lambda}{2}\right\} + \frac{1}{\epsilon}\left\|\max\left\{g(x), -\epsilon\frac{\lambda}{2}\right\}\right\|^2
$$

LANCELOT B

Ultima versione del pacchetto `LANCELOT` disponibile nella collezione di routine per l'ottimizzazione nonlineare `GALAHAD`[2]

[2] `http://www.galahad.rl.ac.uk/`

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

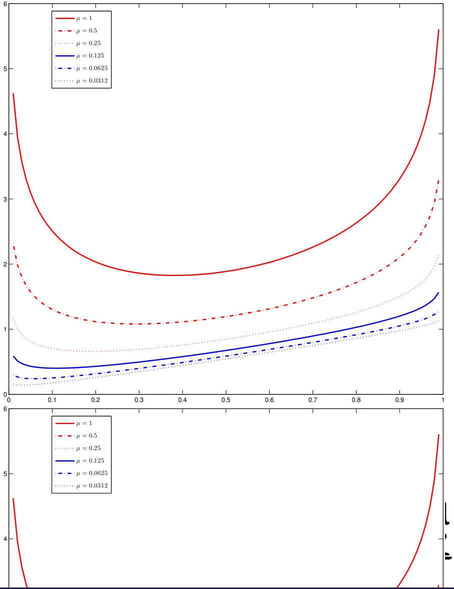
La più diffusa funzione di "barriera" per il problema considerato è:

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

N.B.

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

Esempio 1



$$\begin{aligned} \min_{x} \quad & x \\ c.v. \quad & x \geq 0 \\ & 1 - x \geq 0 \end{aligned}$$

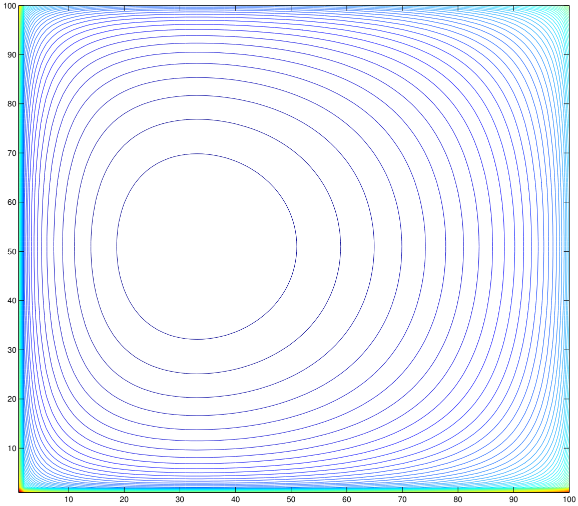
Esempio 2

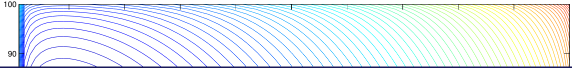
$$\begin{array}{ll} \min\limits_{x} & (x+0.5)^2 + (x_2 - 0.5)^2 \\ c.v. & x_1 \in [0,1],\ x_2 \in [0,1] \end{array}$$

La funzione di barriera per il problema è

$$\begin{aligned} P(x;\mu) = (x+0.5)^2 + (x_2 - 0.5)^2 \\ -\mu[\log x_1 + \log(1-x_1)] \\ -\mu[\log x_2 + \log(1-x_2)] \end{aligned}$$

Esempio 2





Esempio 2

per $\mu = 1 \rightarrow 0.0085$ abbiamo