Ottimizzazione dei Sistemi Complessi

G. Liuzzi[1]

Lunedì 11 Aprile 2019

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

Metodo di soluzione - Metodo dei Moltiplicatori

Algoritmo `SEQLAGR`

Dati: $\epsilon_\circ > 0$, $\beta > 1$, $\{\tau_k\} \to 0$, $\rho > 0$, $\mu_\circ$, `maxit`

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

Calcola x_k t.c. $\|\nabla_x L_a(x_k, \mu_k; \epsilon_k)\| \le \tau_k$

if $\|\nabla L_a(x_k, \mu_k; \epsilon_k)\| < \rho$ **then**

$x^* \leftarrow x_k$, $\mu^* \leftarrow \mu_k$, STOP

endif

$\epsilon_{k+1} = \epsilon_k / \beta$

$$\boxed{\mu_{k+1} = \mu_k + \frac{2h(x_k)}{\epsilon_k}}$$

endfor

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

Proprietà di convergenza

Teorema

Sia $\{x_k\}$ *una successione t.c.* $\|\nabla L_a(x_k, \mu_k; \epsilon_k)\| \leq \tau_k$ *e* $\{x_k\}_K$ *una sottosucc. convergente ad* x^* *t.c. siano lin. indipendenti*

$$\nabla h_j(x^*),\ \forall\, j.$$

Allora, x^* *è un punto di KKT con molt.* μ^* *t.c.*

$$\left\{\mu_k + \frac{2}{\epsilon_k} h(x_k)\right\}_K \to \mu^*$$

Problemi con soli vincoli di disuguaglianza – 1

Consideriamo il problema

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

Usiamo la trasformazione:

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

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

Problemi con soli vincoli di disuguaglianza – 1

Consideriamo il problema

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

Usiamo la trasformazione:

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

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

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

Problemi con soli vincoli di disuguaglianza – 1

Consideriamo il problema

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

Usiamo la trasformazione:

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

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

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

Problemi con soli vincoli di disuguaglianza – 1

Consideriamo il problema

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

Usiamo la trasformazione:

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

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

Problemi con soli vincoli di disuguaglianza – 2

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 – 2

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

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

fissati ϵ e λ, bisogna "risolvere"

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

Problemi con soli vincoli di disuguaglianza – 2

Quindi

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

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 – 2

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"

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

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

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

che è verificata per

$$y_i^2 = 0, \text{ oppure } y_i^2 = \max\left\{0, -\epsilon\frac{\lambda_i}{2} - g_i(x)\right\}$$

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

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

che è verificata per

$$
y_i^2 = 0, \text{ oppure } y_i^2 = \max\left\{0, -\epsilon\frac{\lambda_i}{2} - g_i(x)\right\}
$$

Gradiente della Lagrangiana – segue

Calcoliamo $\nabla^2_{y_i} L_a$

$$\nabla^2_{y_i} L_a = \frac{4}{\epsilon}\left(\epsilon\frac{\lambda_i}{2} + g_i(x) + 3y_i^2\right)$$

- Quando $y_i^2 = 0$, $\nabla^2_{y_i} L_a = \frac{4}{\epsilon}\left(\epsilon\frac{\lambda_i}{2} + g_i(x)\right)$
- Quando $y_i^2 = -\epsilon\frac{\lambda_i}{2} - g_i(x) > 0$, $\nabla^2_{y_i} L_a = \frac{8}{\epsilon}\left(-\epsilon\frac{\lambda_i}{2} - g_i(x)\right)$

Imponendo $\nabla^2_{y_i} L_a > 0$, segue che la soluzione ottima è proprio $y_i^2 = \max\left\{0, -\epsilon\frac{\lambda_i}{2} - g_i(x)\right\}$

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

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

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

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

Metodo di soluzione - Metodo dei Moltiplicatori

Algoritmo `SEQLAGR`

Dati: $\epsilon_\circ > 0$, $\beta > 1$, $\{\tau_k\} \to 0$, $\rho > 0$, $\lambda_\circ$, $\mu_\circ$, `maxit`

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

Calcola x_k t.c. $\|\nabla_x L_a(x_k, \mu_k, \lambda_k; \epsilon_k)\| \leq \tau_k$

if $\|\nabla L_a(x_k, \mu_k, \lambda_k; \epsilon_k)\| < \rho$ **then**

$x^* \leftarrow x_k$, $\mu^* \leftarrow \mu_k$, $\lambda^* \leftarrow \lambda_k$ STOP

endif

$\epsilon_{k+1} = \epsilon_k/\beta$

$$\boxed{\mu_{k+1} = \mu_k + \frac{2h(x_k)}{\epsilon_k},\ \lambda_{k+1} = \lambda_k + \frac{2\max\{g(x_k), -\epsilon_k\lambda_k/2\}}{\epsilon_k}}$$

endfor

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

Proprietà di convergenza

Teorema

Sia $\{x_k\}$ *una successione t.c.* $\|\nabla L_a(x_k,\lambda_k,\mu_k;\epsilon_k)\| \leq \tau_k$ *e* $\{x_k\}_K$ *una sottosucc. convergente ad* x^* *t.c. siano lin. indipendenti*

$$\nabla g_i(x^*),\ \forall\ i \in I_0(x^*),\ \nabla h_j(x^*),\ \forall\ j.$$

Allora, x^* *è un punto di KKT con molt.* λ^* *e* μ^* *t.c.*

$$\left\{\lambda_k + \frac{2}{\epsilon_k}\max\{g(x_k), -\frac{\epsilon_k}{2}\lambda_k\}\right\}_K \to \lambda^*$$

$$\left\{\mu_k + \frac{2}{\epsilon_k}h(x_k)\right\}_K \to \mu^*$$

Un po' di esempi di risoluzione

`esempio3`:

$$\begin{array}{l} \min\ -x_1 - x_2 \\ s.t.\ x_1^2 + x_2^2 = 1 \end{array}$$

$x^0 = (-1,\ -1)^\top$
Ammete $(x^*, y^*) = (1,\ 1)/\sqrt{2}$ come **unico** punto di minimo globale

La funzione Lagrangiana aumentata è:

$$L_a(x, y, \mu; \epsilon) = -x - y + \mu(x_1^2 + x_2^2 - 1) + \frac{1}{\epsilon}(1 - x^2 - y^2)^2$$

Un po' di esempi di risoluzione

`esempio3`:

$$\begin{array}{l} \min \ -x_1 - x_2 \\ s.t. \ x_1^2 + x_2^2 = 1 \end{array}$$

$x^0 = (-1,\ -1)^\top$
Ammete $(x^*, y^*) = (1,\ 1)/\sqrt{2}$ come **unico** punto di minimo globale

La funzione Lagrangiana aumentata è:

$$L_a(x, y, \mu; \epsilon) = -x - y + \mu(x_1^2 + x_2^2 - 1) + \frac{1}{\epsilon}(1 - x^2 - y^2)^2$$

Un po' di esempi di risoluzione

`esempio4`:

$$\begin{array}{l} \min\ -x_1 - x_2 \\ s.t.\ x_1^2 + x_2^2 \leq 1 \end{array}$$

$x^0 = (-1,\ -1)^\top$

Un po' di esempi di risoluzione

`maratos`:

$$\begin{array}{l} \min \ -x_1 + 2(x_1^2 + x_2^2 - 1) \\ s.t. \ x_1^2 + x_2^2 = 1 \end{array}$$

$x^0 = (-1, \ -1)^\top$

Un po' di esempi di risoluzione

`hs14`:

$$\begin{aligned} \min\ & (x_1 - 2)^2 + (x_2 - 1)^2 \\ s.t.\ & x_1^2/4 + x_2^2 \le 1 \\ & x_1 - 2x_2 = -1 \end{aligned}$$

$x^0 = (2,\ 2)^\top$

Un po' di esempi di risoluzione

`hs24`:

$$
\begin{aligned}
\min\ & x_2^3((x_1-3)^2-9)/(27\sqrt{3}) \\
s.t.\ & x_1/\sqrt{3}-x_2 \ge 0 \\
& x_1+\sqrt{3}x_2 \ge 0 \\
& -x_1-\sqrt{3}x_2+6 \ge 0 \\
& x_1 \ge 0,\ \ x_2 \ge 0
\end{aligned}
$$

$x^0 = (1,\ 1/2)^\top$

Un po' di esempi di risoluzione

`hs32`:

$$\begin{aligned} &\min\ (x_1 + 3 * x_2 + x_3)^2 + 4(x_1 - x_2)^2 \\ &s.t.\ x_1 + x_2 + x_3 = 1 \\ &\qquad 6x_2 + 4x_3 - x_1^3 \geq 3 \\ &\qquad x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0 \end{aligned}$$

$x^0 = (0.1,\ 0.7,\ 0.2)^\top$

Un po' di esempi di risoluzione

`hs41`:

$$
\begin{aligned}
\min\ & 2 - x_1x_2x_3 \\
s.t.\ & x_1 + 2x_2 + 2x_3 - x_4 = 0 \\
& 0 \le x_i \le 1,\ \ i = 1, 2, 3, 4
\end{aligned}
$$

$x^0 = (2,\ 2,\ 2,\ 2)^\top$

Un po' di esempi di risoluzione

`hs55`:

$$
\begin{aligned}
\min\ & x_1 + 2x_2 + 4x_5 + e^{x_1 x_4} \\
s.t.\ & x_1 + 2x_2 + 5x_5 - 6 = 0 \\
& x_1 + x_2 + x_3 - 3 = 0 \\
& x_4 + x_5 + x_6 - 2 = 0 \\
& x_1 + x_4 - 1 = 0 \\
& x_2 + x_5 - 2 = 0 \\
& x_3 + x_6 - 2 = 0 \\
& x_i \geq 0,\ i = 1, \dots, 6 \\
& x_1 \leq 1,\ x_4 \leq 1
\end{aligned}
$$

$x^0 = (1,\ 2,\ 0,\ 0,\ 0,\ 2)^\top$

Un po' di esempi di risoluzione

`hs60`:

$$\begin{aligned} &\min\ (x_1-1)^2+(x_1-x_2)^2+(x_2-x_3)^4 \\ &s.t.\ x_1(1+x_2^2)+x_3^4-4-3\sqrt{2}=0 \\ &\qquad -10 \le x_i \le 10,\ \ i=1,2,3 \end{aligned}$$

$x^0 = (2,\ 2,\ 2)^\top$