Ottimizzazione Non Lineare

G. Liuzzi[1]

Giovedì 6 Dicembre 2018

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

Definizione del problema

Il metodo DIRECT considera il problema

$$\begin{aligned} &\min\ f(x) \\ &s.t.\ \mathbf{0} \le x \le \mathbf{1} \end{aligned} \tag{1}$$

per cui $\mathcal{D} = \{x \in \mathbb{R}^n : \ \mathbf{0} \le x \le \mathbf{1}\}$

Nota: qualunque problema

$$\begin{aligned} &\min\ f(x) \\ &s.t.\ \ell \le x \le u \end{aligned}$$

con $-\infty < \ell_i < u_i < +\infty$, $i = 1, \dots, n$, può essere riportato nella forma del problema (1)

Definizione del problema

Il metodo `DIRECT` considera il problema

$$\begin{array}{l} \min \ f(x) \\ s.t.\ \mathbf{0} \le x \le \mathbf{1} \end{array} \qquad (1)$$

per cui $\mathcal{D} = \{x \in \mathbb{R}^n : \ \mathbf{0} \le x \le \mathbf{1}\}$

Nota: qualunque problema

$$\begin{array}{l} \min \ f(x) \\ s.t.\ \ell \le x \le u \end{array}$$

con $-\infty < \ell_i < u_i < +\infty$, $i = 1, \dots, n$, può essere riportato nella forma del problema (1)

Un metodo di partizionamento

DIRECT produce, iterazione dopo iterazione, partizioni sempre più fini dell'insieme ammissibile iniziale $\mathcal{D}$.

Se $\mathcal{H}_k = \{\mathcal{D}^i : \ i \in I_k\}$ è la partizione all'inizio della iterazione k

la partizione all'inizio della iterazione $k+1$ sarà
$\mathcal{H}_{k+1} = \{\mathcal{D}^i : \ i \in I_{k+1}\}$ con

$$|\mathcal{H}_{k+1}| > |\mathcal{H}_k|$$

Schema generale di DIRECT

Set $\mathcal{D}^0 = \mathcal{D}$, $\mathcal{H}_0 = \{D^0\}$, $I_0 = \{0\}$, $c = \mathbf{1}/2$, $k = 0$

repeat

Usa una $\boxed{\text{procedura di selezione}}$ per determinare $I_k^* \subseteq I_k$

Usa una $\boxed{\text{procedura di partizione}}$ per dividere $\mathcal{D}^i$, $i \in I_k^*$

Definisci la nuova partizione $\mathcal{H}_{k+1}$

Set $k = k + 1$

until (stopping criterion satisfied)

return $f_{min} = \min\{f(c) : c \in C_k\}$, $X_{min} = \{c \in C_k : f(c) = f_{min}\}$,
with $C_k = \{\text{centri degli insiemi in } \mathcal{H}_k\}$

Procedura di selezione

Definizione

Data la partizione $\{\mathcal{D}^i : i \in I\}$ *con* $\mathcal{D}^i = \{x : \; l_i \leq x \leq u_i\}$ *e* $x^i = \frac{u_i + l_i}{2}$.
$\mathcal{D}^h$ *è* **potenzialmente ottimo** *se, scelto un* $\epsilon > 0$, *esiste una costante* $L^h > 0$ *tale che*

$$\begin{aligned} f(x^h) - L^h \frac{\|u^h - l^h\|}{2} &\leq f(x^i) - L^h \frac{\|u^i - l^i\|}{2}, \quad \forall i \in I \\ f(x^h) - L^h \frac{\|u^h - l^h\|}{2} &\leq f_{min} - \epsilon |f_{min}| \end{aligned}$$

$$I_k^* = \{i \in I_k : \mathcal{D}^i \text{ è potenzialmente ottimo}\}$$

Procedura di selezione

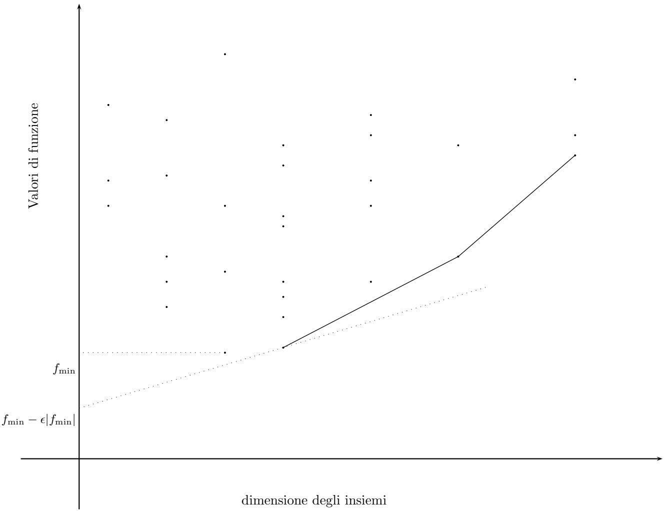
Definizione

Data la partizione $\{\mathcal{D}^i : i \in I\}$ *con* $\mathcal{D}^i = \{x : \; l_i \le x \le u_i\}$ *e* $x^i = \frac{u_i + l_i}{2}$.
$\mathcal{D}^h$ *è* **potenzialmente ottimo** *se, scelto un* $\epsilon > 0$, *esiste una costante* $L^h > 0$ *tale che*

$$
\begin{aligned}
f(x^h) - L^h \frac{\|u^h - l^h\|}{2} &\le f(x^i) - L^h \frac{\|u^i - l^i\|}{2}, \quad \forall i \in I \\
f(x^h) - L^h \frac{\|u^h - l^h\|}{2} &\le f_{min} - \epsilon |f_{min}|
\end{aligned}
$$

$$I_k^* = \{i \in I_k : \mathcal{D}^i \text{ è potenzialmente ottimo}\}$$

Procedura di selezione



Procedura di partizione

Supponiamo che $h \in I_k^*$ e cioè che si debba ulteriormente partizionare $\mathcal{D}^h$. Siano

$$
\begin{aligned}
\delta &= \max_{1 \le j \le n} (u^h - \ell^h)_j, \\
J &= \{ j \in \{1, \dots, n\} : \ (u^h - \ell^h)_j = \delta \}, \\
m &= |J|
\end{aligned}
$$

Per ogni $j \in J$, si valuta la funzione obiettivo sui punti

$$
c^{h_j} = c^h + \frac{\delta}{3} e_j, \qquad c^{h_{j+m}} = c^h - \frac{\delta}{3} e_j,
$$

Procedura di partizione

Supponiamo che $h \in I_k^*$ e cioè che si debba ulteriormente partizionare $\mathcal{D}^h$. Siano

$$\begin{aligned} \delta &= \max_{1 \le j \le n} (u^h - \ell^h)_j, \\ J &= \{j \in \{1, \ldots, n\} : \ (u^h - \ell^h)_j = \delta\}, \\ m &= |J| \end{aligned}$$

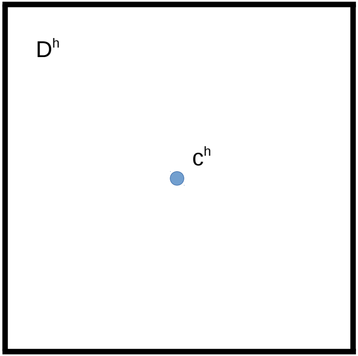
Per ogni $j \in J$, si valuta la funzione obiettivo sui punti

$$c^{h_j} = c^h + \frac{\delta}{3} e_j, \qquad c^{h_{j+m}} = c^h - \frac{\delta}{3} e_j,$$

Procedura di partizione

L'insieme $\mathcal{D}^h$ viene partizionato in $2m+1$ sottoinsiemi tali che:

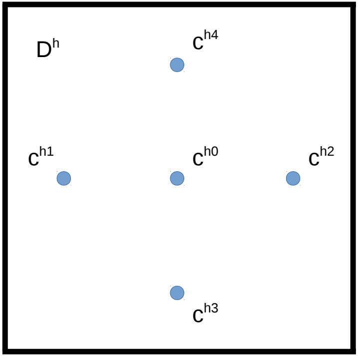
- hanno i punti c^h, c^{h_j}, $c^{h_{j+m}}$, $j \in J$, come centri;
- se $m > 1$, si fa in modo che gli insiemi più grandi contengono i punti con il miglior valore di funzione obiettivo



Procedura di partizione

L'insieme $\mathcal{D}^h$ viene partizionato in $2m+1$ sottoinsiemi tali che:

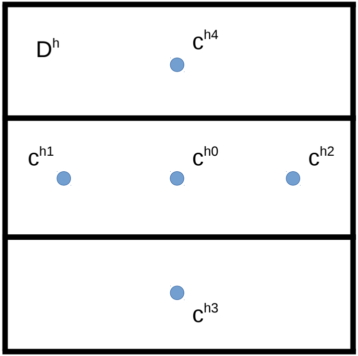
- hanno i punti c^h, c^{h_j}, $c^{h_{j+m}}$, $j \in J$, come centri;
- se $m > 1$, si fa in modo che gli insiemi più grandi contengono i punti con il miglior valore di funzione obiettivo



Procedura di partizione

L'insieme $\mathcal{D}^h$ viene partizionato in $2m+1$ sottoinsiemi tali che:

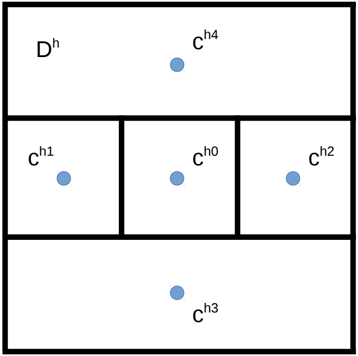
- hanno i punti c^h, c^{h_j}, $c^{h_{j+m}}$, $j \in J$, come centri;
- se $m > 1$, si fa in modo che gli insiemi più grandi contengono i punti con il miglior valore di funzione obiettivo



Procedura di partizione

L'insieme $\mathcal{D}^h$ viene partizionato in $2m+1$ sottoinsiemi tali che:

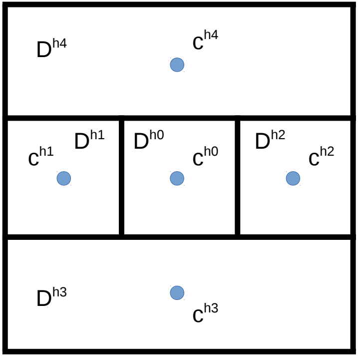
- hanno i punti c^h, c^{h_j}, $c^{h_{j+m}}$, $j \in J$, come centri;
- se $m > 1$, si fa in modo che gli insiemi più grandi contengono i punti con il miglior valore di funzione obiettivo



Procedura di partizione

L'insieme $\mathcal{D}^h$ viene partizionato in $2m+1$ sottoinsiemi tali che:

- hanno i punti c^h, c^{h_j}, $c^{h_{j+m}}$, $j \in J$, come centri;
- se $m > 1$, si fa in modo che gli insiemi più grandi contengono i punti con il miglior valore di funzione obiettivo



Esempio (1)

Supponiamo $n = 2$ e sia $\mathcal{D}^h$ l'insieme iniziale $\mathcal{D}$, cioè l'iperbox unitario.

In questo caso:

$$\begin{aligned}
\delta &= \max_{1 \le j \le n} (u^h - \ell^h)_j = 1, \\
J &= \{j \in \{1, \dots, n\} : \ (u^h - \ell^h)_j = 1\} = \{1, 2\}, \\
m &= |J| = 2
\end{aligned}$$

Valutiamo la funzione sui punti

$$\begin{aligned}
c^{h_0} &= c^h = \{1/2, 1/2\} \\
c^{h_1} &= c^h + e_1/3 = \{5/6, 1/2\}, \quad c^{h_3} = c^h - e_1/3 = \{1/6, 1/2\} \\
c^{h_2} &= c^h + e_2/3 = \{1/2, 5/6\}, \quad c^{h_4} = c^h - e_2/3 = \{1/2, 1/6\}
\end{aligned}$$

Esempio (1)

Supponiamo che risulti

$$\begin{aligned} f(c^{h_1}) &= 3, \quad f(c^{h_3}) = 7 \\ f(c^{h_2}) &= 5, \quad f(c^{h_4}) = 6 \end{aligned}$$

Disegnare la partizione di $\mathcal{D}^h$ così come determinata dalla procedura di partizione.

Esempio (2)

Supponiamo $n = 2$ e sia $\mathcal{D}^h$ l'insieme

$$\mathcal{D}^h = \{x \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, 1/3 \leq x_2 \leq 2/3\}$$

In questo caso:

$$\begin{aligned} \delta &= \max_{1 \leq j \leq n} (u^h - \ell^h)_j = 1, \\ J &= \{j \in \{1, \ldots, n\} : \; (u^h - \ell^h)_j = 1\} = \{1\}, \\ m &= |J| = 1 \end{aligned}$$

Valutiamo la funzione sui punti

$$\begin{aligned} c^{h_0} &= c^h = \{1/2, 1/2\} \\ c^{h_1} &= c^h + e_1/3 = \{5/6, 1/2\}, \quad c^{h_2} = c^h - e_1/3 = \{1/6, 1/2\} \end{aligned}$$

Esempio (2)

Disegnare la partizione di $\mathcal{D}^h$ così come determinata dalla procedura di partizione.

Riferimenti

1. F. Schoen, M. Locatelli, “Global Optimization: Theory, Algorithms, and Applications” MOS-SIAM series on Optimization, SIAM Philadelphia
2. S. Lucidi, “Appunti dalle lezioni di OTTIMIZZAZIONE GLOBALE” http://www.diag.uniroma1.it/∼lucidi/didattica/Ott-Glob.html
3. http://www.iasi.cnr.it/∼liuzzi/DFL