Ottimizzazione dei Sistemi Complessi

G. Liuzzi[1]

Giovedì 12 Marzo 2020

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

Iterazione k di Nelder& Mead

- Se $\boxed{f_1 \leq f^r < f_n}$, allora $X_{k+1} = X_k \setminus \{x_{n+1}\} \cup \{x^r\}$ **FINE**
- Se $\boxed{f^r < f_1}$, allora
  - Se $f^e < f^r$, allora $X_{k+1} = X_k \setminus \{x_{n+1}\} \cup \{x^e\}$ **FINE**
  - altrimenti $X_{k+1} = X_k \setminus \{x_{n+1}\} \cup \{x^r\}$ **FINE**
- Se $\boxed{f_n \leq f^r < f_{n+1}}$, allora
  - Se $f^{oc} \leq f^r$, allora $X_{k+1} = X_k \setminus \{x_{n+1}\} \cup \{x^{oc}\}$ **FINE**
  - altrimenti **shrink**
- Se $\boxed{f_{n+1} \leq f^r}$, allora
  - Se $f^{ic} < f_{n+1}$, allora $X_{k+1} = X_k \setminus \{x_{n+1}\} \cup \{x^{ic}\}$ **FINE**
  - altrimenti **shrink**
- **shrink**:
  - $X_{k+1} = \{x_1, \hat{x}_2, \ldots, \hat{x}_{n+1}\}$ dove
  - $\hat{x}_i = x_1 + \gamma(x_i - x_1)$, $i = 2, \ldots, n+1$, $\gamma \in (0,1)$

Esercizio

Sia

$$
\begin{array}{rcl}
X & = & \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} \\
F & = & \{ \quad 3, \qquad 7, \qquad 1, \quad \}
\end{array}
$$

Il centroide $\bar{x}$ di x_1 e x_2 (dopo il riordino) è

$$\bar{x} = \begin{pmatrix} 1 \\ 1/2 \end{pmatrix}$$

Il metodo considera i punti

$$x(\mu) = \bar{x} + \mu d = \bar{x} + \mu(\bar{x} - x_3)$$

per valori $\mu = 1,\ 2,\ 1/2,\ -1/2$

Esercizio

Sia

$$
\begin{aligned}
X &= \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \\
F &= \{ \quad 1, \qquad 3, \qquad 7, \quad \}
\end{aligned}
$$

Il centroide $\bar{x}$ di x_1 e x_2 (dopo il riordino) è

$$\bar{x} = \begin{pmatrix} 1 \\ 1/2 \end{pmatrix}$$

Il metodo considera i punti

$$x(\mu) = \bar{x} + \mu d = \bar{x} + \mu(\bar{x} - x_3)$$

per valori $\mu = 1,\ 2,\ 1/2,\ -1/2$

Esercizio

Sia

$$X = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$
$$F = \{ \quad 1, \quad 3, \quad 7, \quad \}$$

Il centroide $\bar{x}$ di x_1 e x_2 (dopo il riordino) è

$$\bar{x} = \begin{pmatrix} 1 \\ 1/2 \end{pmatrix}$$

Il metodo considera i punti

$$x(\mu) = \bar{x} + \mu d = \bar{x} + \mu(\bar{x} - x_3)$$

per valori $\mu = 1,\ 2,\ 1/2,\ -1/2$

Esercizio

Sia

$$\begin{aligned} X &= \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \\ F &= \{ \quad 1, \qquad 3, \qquad 7, \quad \} \end{aligned}$$

Il centroide $\bar{x}$ di x_1 e x_2 (dopo il riordino) è

$$\bar{x} = \begin{pmatrix} 1 \\ 1/2 \end{pmatrix}$$

Il metodo considera i punti

$$x(\mu) = \bar{x} + \mu d = \bar{x} + \mu(\bar{x} - x_3)$$

per valori $\mu = 1,\ 2,\ 1/2,\ -1/2$

Esercizio

- si calcola $d = \bar{x} - x_3 = (1,\ -1/2)^\top$
- si calcola $x_r = x(1) = \bar{x} + d = (2,\ 0)^\top$
- si calcola $x_e = x(2) = \bar{x} + 2d = (3,\ -1/2)^\top$
- si calcola $x_{oc} = x(1/2) = \bar{x} + 1/2d = (3/2,\ 1/4)^\top$
- si calcola $x_{ic} = x(-1/2) = \bar{x} - 1/2d = (1/2,\ 3/4)^\top$

Esercizio

- si calcola $d = \bar{x} - x_3 = (1,\ -1/2)^\top$
- si calcola $x_r = x(1) = \bar{x} + d = (2,\ 0)^\top$
- si calcola $x_e = x(2) = \bar{x} + 2d = (3,\ -1/2)^\top$
- si calcola $x_{oc} = x(1/2) = \bar{x} + 1/2d = (3/2,\ 1/4)^\top$
- si calcola $x_{ic} = x(-1/2) = \bar{x} - 1/2d = (1/2,\ 3/4)^\top$

Esercizio

- si calcola $d = \bar{x} - x_3 = (1,\ -1/2)^\top$
- si calcola $x_r = x(1) = \bar{x} + d = (2,\ 0)^\top$
- si calcola $x_e = x(2) = \bar{x} + 2d = (3,\ -1/2)^\top$
- si calcola $x_{oc} = x(1/2) = \bar{x} + 1/2d = (3/2,\ 1/4)^\top$
- si calcola $x_{ic} = x(-1/2) = \bar{x} - 1/2d = (1/2,\ 3/4)^\top$

Esercizio

- si calcola $d = \bar{x} - x_3 = (1,\ -1/2)^\top$
- si calcola $x_r = x(1) = \bar{x} + d = (2,\ 0)^\top$
- si calcola $x_e = x(2) = \bar{x} + 2d = (3,\ -1/2)^\top$
- si calcola $x_{oc} = x(1/2) = \bar{x} + 1/2d = (3/2,\ 1/4)^\top$
- si calcola $x_{ic} = x(-1/2) = \bar{x} - 1/2d = (1/2,\ 3/4)^\top$

Esercizio

- si calcola $d = \bar{x} - x_3 = (1,\ -1/2)^\top$
- si calcola $x_r = x(1) = \bar{x} + d = (2,\ 0)^\top$
- si calcola $x_e = x(2) = \bar{x} + 2d = (3,\ -1/2)^\top$
- si calcola $x_{oc} = x(1/2) = \bar{x} + 1/2d = (3/2,\ 1/4)^\top$
- si calcola $x_{ic} = x(-1/2) = \bar{x} - 1/2d = (1/2,\ 3/4)^\top$

Nelder&Mead (1)

In $\mathbb{R}^2$ siano dati i seguenti punti:

$$x_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, x_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, x_3 = \begin{pmatrix} 1 \\ 3 \end{pmatrix},$$

a cui corrispondono: $f(x_1) = 1$, $f(x_2) = 3$, $f(x_3) = 5$.

Determinare i punti x_r, x_e, x_{oc} e x_{ic} utilizzati nel metodo di Nelder&Mead

Nelder&Mead (1)

In $\mathbb{R}^2$ siano dati i seguenti punti:

$$x_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, x_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, x_3 = \begin{pmatrix} 1 \\ 3 \end{pmatrix},$$

a cui corrispondono: $f(x_1) = 1$, $f(x_2) = 3$, $f(x_3) = 5$.

Determinare i punti x_r, x_e, x_{oc} e x_{ic} utilizzati nel metodo di Nelder&Mead

Soluzione

1) Calcola $\bar{x} = (2,\ 3/2)^\top$
2) Calcola la direzione $d = \bar{x} - x_3 = (1,\ -3/2)^\top$
3) Calcola i 4 punti richiesti:

$$\begin{aligned} x_r &= \bar{x} + d = (3,\ 0)^\top \\ x_e &= \bar{x} + 2d = (4,\ -3/2)^\top \\ x_{oc} &= \bar{x} + d/2 = (5/2,\ 3/4)^\top \\ x_{ic} &= \bar{x} - d/2 = (3/2,\ 9/4)^\top \end{aligned}$$

Nelder&Mead (2)

In $\mathbb{R}^3$ è dato il seguente insieme X_k di punti

$$X_k = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} \right\}$$

Assegnare, ove necessario, valori di funzione obiettivo per fare in modo che l'iterazione corrente si concluda con una operazione di **SHRINK** dopo aver provato il punto di "**outer contraction**" x_{oc}.

Soluzione

Supponiamo che sia

$$X_k = \{x_1, x_2, x_3, x_4\}, \quad \text{con} \quad f(x_i) = i, \quad i = 1, 2, 3, 4,$$

in modo che i punti nell'insieme X_k risultano già essere ordinati dal "migliore" (valore di f più basso) al "peggiore" (valore di f più alto)

Per tanto, l'algoritmo calcolerà $\bar{x} = (2/3,\ 0,\ 2/3)^\top$ e $d = \bar{x} - x_4 = (2/3,\ 1,\ 2/3)^\top$. Ora immaginiamo che sia

$f(x_r) = 3.5$ con $x_r = \bar{x} + d = (4/3,\ 1,\ 4/3)^\top$ e

$f(x_{oc}) = 4$ con $x_{oc} = \bar{x} + d/2 = (1,\ 1/2,\ 1)^\top$

allora verrà eseguita l'operazione di shrink a fine iterazione.

Esercizio

Si consideri il problema

$$\min_{x \in \mathbb{R}^n} \; f(x)$$

con

- $f(x)$ continuamente differenziabile;
- $\nabla f(x)$ Lipschitz continuo modulo $L = 10$.

Sia inoltre $x_0 \in \mathbb{R}^n$ tale che $L_0 = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$ è compatto.

Supponiamo che, per la soluzione del problema, si adotti un metodo tipo "compass search" con $\Delta_0 = 1$ tale che $\Delta_{k+1} \leq \Delta_k$. In particolare $\Delta_{k+1} = \Delta_k/2$ se e soltanto se la k-esima iterazione è di "fallimento".

Stabilire quale è il numero minimo di iterazioni di fallimento che si devono avere per poter affermare con certezza che, se la k-esima iterazione è una di fallimento, si avrà:

$$\|\nabla f(x_k)\| \leq 10^{-6}$$

Esercizio – soluzione

Ci ricordiamo che nelle iterazioni di fallimento (proprio quelle cioè in cui risulta $\Delta_{k+1} = \Delta_k/2 < \Delta_k$) sappiamo dare una limitazione superiore alla norma del gradiente, ovvero:

$$\|\nabla f(x_k)\| \leq L\sqrt{n}\Delta_k.$$

Se quindi facciamo in modo che, nelle iterazioni di fallimento, risulti

$$L\sqrt{n}\Delta_k \leq 10^{-6}$$

potremo essere sicuri che $\|\nabla f(x_k)\| \leq 10^{-6}$. La domanda è ora quindi, quante iterazioni di fallimento (fino alla k-esima) sono necessarie affinché risulti

$$\Delta_k \leq \frac{10^{-6}}{L\sqrt{n}}$$

Se indichiamo con ℓ il numero di iterazioni di fallimento fino alla iterazione k (di fallimento anche essa) abbiamo che $\Delta_k = \Delta_0/2^\ell$ pertanto

$$2^\ell \geq \frac{\Delta_0 L\sqrt{n}}{10^{-6}}$$

cioè $\ell \geq \left\lceil \log_2\left(\frac{\Delta_0 L\sqrt{n}}{10^{-6}}\right)\right\rceil$ iterazioni di fallimento.

Funzione di Broyden