Ottimizzazione dei Sistemi Complessi

G. Liuzzi[1]

Mercoledì 4 Marzo 2020

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

Pseudo-codice del "compass search" forte

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $D = \{\pm e_i, i = 1, \ldots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq \mathtt{maxit}$ **and** $\Delta_k \geq \Delta_{min}$ **do**

> $k \leftarrow k + 1$
>
> Let $\bar{d} \in D$ be s.t. $f(x_k + \Delta_k \bar{d}) = \min_{d_i \in D} f(x_k + \Delta_k d_i)$
>
> **if** $f(x_k + \Delta_k \bar{d}) < f(x_k)$ **then**
>
> > $x_{k+1} \leftarrow x_k + \Delta_k \bar{d}$, $\Delta_{k+1} \leftarrow \Delta_k$
>
> **else**
>
> > $\Delta_{k+1} \leftarrow \Delta_k / 2$, $x_{k+1} \leftarrow x_k$
>
> **endif**
>
> $k \leftarrow k + 1$

end while

RETURN: x (miglior punto determinato)

Pseudo-codice del "compass search" forte

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $D = \{\pm e_i, i = 1, \dots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq \mathtt{maxit}$ **and** $\Delta_k \geq \Delta_{min}$ **do**

$k \leftarrow k + 1$

Let $\bar{d} \in D$ be s.t. $f(x_k + \Delta_k \bar{d}) = \min_{d_i \in D} f(x_k + \Delta_k d_i)$

if $f(x_k + \Delta_k \bar{d}) < f(x_k)$ **then**

$x_{k+1} \leftarrow x_k + \Delta_k \bar{d}$, $\Delta_{k+1} \leftarrow \Delta_k$

else

$\Delta_{k+1} \leftarrow \Delta_k / 2$, $x_{k+1} \leftarrow x_k$

endif

$k \leftarrow k + 1$

end while

RETURN: $\{x_k\}, \{\Delta_k\}$ successioni di punti e passi

Un po' di analisi (matematica)

Teorema (della media o di Lagrange in $\mathbb{R}$)

$f(x) : \mathbb{R} \to \mathbb{R}$ *continua e derivabile in* $[a, b]$. *Esiste sempre un* $\xi \in (a, b)$ *tale che*

$$f(b) - f(a) = f'(\xi)(b - a)$$

Definizione (Segmento in $\mathbb{R}^n$)

Dati due punti $a, b \in \mathbb{R}^n$, *il segmento di estremi* a *e* b *è*

$$[a, b] \triangleq \{x \in \mathbb{R}^n : x = ta + (1-t)b,\ t \in [0, 1]\}$$
$$(a, b) \triangleq \{x \in \mathbb{R}^n : x = ta + (1-t)b,\ t \in (0, 1)\}$$

Teorema (della media o di Lagrange in $\mathbb{R}^n$)

$f(x) : \mathbb{R}^n \to \mathbb{R}$ *continua e derivabile in* $[a, b]$. *Esiste sempre un* $\xi \in (a, b)$ *tale che*

$$f(b) - f(a) = \nabla f(\xi)^\top (b - a)$$

Un po' di analisi (matematica)

Teorema (della media o di Lagrange in $\mathbb{R}$)

$f(x) : \mathbb{R} \to \mathbb{R}$ continua e derivabile in $[a, b]$. Esiste sempre un $\xi \in (a, b)$ tale che

$$f(b) - f(a) = f'(\xi)(b - a)$$

Definizione (Segmento in $\mathbb{R}^n$)

Dati due punti $a, b \in \mathbb{R}^n$, il segmento di estremi a e b è

$$\begin{aligned} [a, b] &\triangleq \{x \in \mathbb{R}^n : x = ta + (1-t)b,\ t \in [0, 1]\} \\ (a, b) &\triangleq \{x \in \mathbb{R}^n : x = ta + (1-t)b,\ t \in (0, 1)\} \end{aligned}$$

Teorema (della media o di Lagrange in $\mathbb{R}^n$)

$f(x) : \mathbb{R}^n \to \mathbb{R}$ continua e derivabile in $[a, b]$. Esiste sempre un $\xi \in (a, b)$ tale che

$$f(b) - f(a) = \nabla f(\xi)^\top (b - a)$$

Un po' di analisi (matematica)

Teorema (della media o di Lagrange in $\mathbb{R}$)

$f(x) : \mathbb{R} \to \mathbb{R}$ continua e derivabile in $[a, b]$. Esiste sempre un $\xi \in (a, b)$ tale che

$$f(b) - f(a) = f'(\xi)(b - a)$$

Definizione (Segmento in $\mathbb{R}^n$)

Dati due punti $a, b \in \mathbb{R}^n$, il segmento di estremi a e b è

$$\begin{aligned} [a, b] &\triangleq \{x \in \mathbb{R}^n : x = ta + (1 - t)b,\ \ t \in [0, 1]\} \\ (a, b) &\triangleq \{x \in \mathbb{R}^n : x = ta + (1 - t)b,\ \ t \in (0, 1)\} \end{aligned}$$

Teorema (della media o di Lagrange in $\mathbb{R}^n$)

$f(x) : \mathbb{R}^n \to \mathbb{R}$ continua e derivabile in $[a, b]$. Esiste sempre un $\xi \in (a, b)$ tale che

$$f(b) - f(a) = \nabla f(\xi)^\top (b - a)$$

Convergenza a punti stazionari

Sotto l'ipotesi (A1), faremo vedere che

$$\liminf_{k\to\infty} \|\nabla f(x_k)\| = 0.$$

Seguiremo due strade.

1. Assumendo che ∇f sia continuo;
2. Assumendo che ∇f sia Lipschitz continuo con costante L.

Lipschitz Continuità di ∇f

Definizione

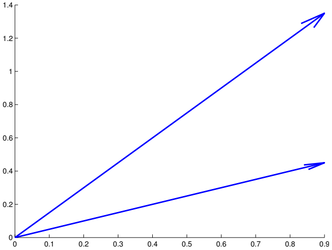
∇f è Lipschitz continuo (con costante L) su un insieme $A \subseteq \mathbb{R}^n$ quando, comunque presi due punti $x, y \in A$, risulta

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|.$$

Angolo tra due vettori

Coseno dell'angolo tra due vettori in $\mathbb{R}^n$:

$$\cos\theta(d,v) = \frac{v^\top d}{\|v\|\|d\|}$$

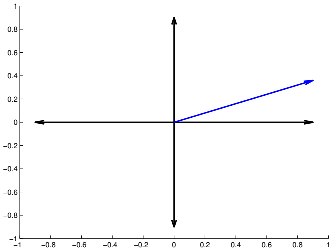


Angolo vettore – insieme di vettori

Dati:

- un insieme finito di vettori (direzioni) D
- un vettore $v \in \mathbb{R}^n$

$$\phi(v) = \max_{d \in D} \frac{v^\top d}{\|v\| \|d\|} = \max_{d \in D} \cos \theta(d, v)$$



Misura del coseno (di un ins. di vettori)

$$\phi(v) = \max_{d \in D} \frac{v^\top d}{\|v\|\|d\|} = \max_{d \in D} \cos\theta(d, v)$$

Dato un insieme finito di vettori (direzioni) D

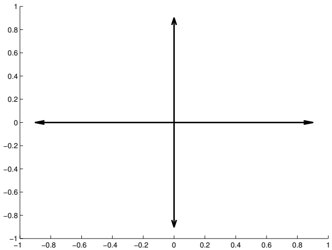
$$\kappa(D) = \min_{v \in \mathbb{R}^n} \phi(v) = \min_{v \in \mathbb{R}^n} \max_{d \in D} \frac{v^\top d}{\|v\|\|d\|}$$

Misura del coseno (di un ins. di vettori)

$$\phi(v) = \max_{d \in D} \frac{v^\top d}{\|v\|\|d\|} = \max_{d \in D} \cos\theta(d, v)$$

Dato un insieme finito di vettori (direzioni) D

$$\kappa(D) = \min_{v \in \mathbb{R}^n} \phi(v) = \min_{v \in \mathbb{R}^n} \max_{d \in D} \frac{v^\top d}{\|v\|\|d\|}$$



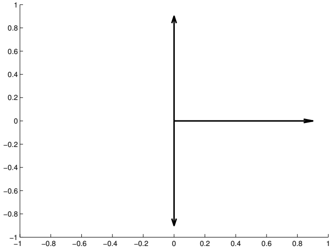
$$\kappa(D) = \frac{1}{\sqrt{2}}$$

Misura del coseno (di un ins. di vettori)

$$\phi(v) = \max_{d\in D} \frac{v^\top d}{\|v\|\|d\|} = \max_{d\in D} \cos\theta(d,v)$$

Dato un insieme finito di vettori (direzioni) D

$$\kappa(D) = \min_{v\in\mathbb{R}^n} \phi(v) = \min_{v\in\mathbb{R}^n} \max_{d\in D} \frac{v^\top d}{\|v\|\|d\|}$$



$$\kappa(D) = 0$$

Una proprietà importante

Limitiamoci a considerare $D = \{\pm e_1, \ldots, \pm e_n\}$. Allora

$$\kappa(D) = \min_{v \in \mathbb{R}^n} \max_{d \in D} \frac{v^\top d}{\|v\| \|d\|} = \min_{v \in \mathbb{R}^n} \max_{i=1,\ldots,n} \frac{|v^\top e_i|}{\|v\|}$$

$$= \min_{v \in \mathbb{R}^n, \|v\|=1} \max_{i=1,\ldots,n} |v^\top e_i|$$

Lemma

$\kappa(D) = 1/\sqrt{n}$

dim. Per ogni $v \in \mathbb{R}^n$, $|v^\top e_i| = |v_i|$. Se $\|v\|^2 = \sum_{i=1}^n |v_i|^2 = 1$, allora $|v^\top e_j| \geq 1/\sqrt{n}$ per qualche j. Quindi, $\kappa(D) \geq 1/\sqrt{n}$. Se consideriamo $v = \alpha_1 e_1 + \cdots + \alpha_n e_n$ con $\alpha_i = \pm 1/\sqrt{n}$, $\kappa(D) = 1/\sqrt{n}$. $\square$

Una proprietà importante

Limitiamoci a considerare $D = \{\pm e_1, \ldots, \pm e_n\}$. Allora

$$\kappa(D) = \min_{v\in\mathbb{R}^n} \max_{d\in D} \frac{v^\top d}{\|v\|\|d\|} = \min_{v\in\mathbb{R}^n} \max_{i=1,\ldots,n} \frac{|v^\top e_i|}{\|v\|}$$

$$= \min_{v\in\mathbb{R}^n,\|v\|=1} \max_{i=1,\ldots,n} |v^\top e_i|$$

Lemma

$\kappa(D) = 1/\sqrt{n}$

dim. Per ogni $v \in \mathbb{R}^n$, $|v^\top e_i| = |v_i|$. Se $\|v\|^2 = \sum_{i=1}^n |v_i|^2 = 1$, allora $|v^\top e_j| \geq 1/\sqrt{n}$ per qualche j. Quindi, $\kappa(D) \geq 1/\sqrt{n}$. Se consideriamo $v = \alpha_1 e_1 + \cdots + \alpha_n e_n$ con $\alpha_i = \pm 1/\sqrt{n}$, $\kappa(D) = 1/\sqrt{n}$. $\square$

Una proprietà importante

Limitiamoci a considerare $D = \{\pm e_1, \dots, \pm e_n\}$. Allora

$$\begin{aligned}\kappa(D) &= \min_{v\in\mathbb{R}^n}\max_{d\in D}\frac{v^\top d}{\|v\|\|d\|} = \min_{v\in\mathbb{R}^n}\max_{i=1,\dots,n}\frac{|v^\top e_i|}{\|v\|}\\ &= \min_{v\in\mathbb{R}^n,\|v\|=1}\max_{i=1,\dots,n}|v^\top e_i|\end{aligned}$$

Lemma

$\kappa(D) = 1/\sqrt{n}$

dim. Per ogni $v\in\mathbb{R}^n$, $|v^\top e_i| = |v_i|$. Se $\|v\|^2 = \sum_{i=1}^n |v_i|^2 = 1$, allora $|v^\top e_j| \geq 1/\sqrt{n}$ per qualche j. Quindi, $\kappa(D) \geq 1/\sqrt{n}$. Se consideriamo $v = \alpha_1 e_1 + \cdots + \alpha_n e_n$ con $\alpha_i = \pm 1/\sqrt{n}$, $\kappa(D) = 1/\sqrt{n}$. $\square$

Una proprietà importante

Limitiamoci a considerare $D = \{\pm e_1, \dots, \pm e_n\}$. Allora

$$\begin{aligned}\kappa(D) &= \min_{v\in\mathbb{R}^n}\max_{d\in D}\frac{v^\top d}{\|v\|\|d\|} = \min_{v\in\mathbb{R}^n}\max_{i=1,\dots,n}\frac{|v^\top e_i|}{\|v\|}\\ &= \min_{v\in\mathbb{R}^n,\|v\|=1}\max_{i=1,\dots,n}|v^\top e_i|\end{aligned}$$

Lemma

$\kappa(D) = 1/\sqrt{n}$

dim. Per ogni $v \in \mathbb{R}^n$, $|v^\top e_i| = |v_i|$. Se $\|v\|^2 = \sum_{i=1}^n |v_i|^2 = 1$, allora $|v^\top e_j| \geq 1/\sqrt{n}$ per qualche j. Quindi, $\kappa(D) \geq 1/\sqrt{n}$. Se consideriamo $v = \alpha_1 e_1 + \dots + \alpha_n e_n$ con $\alpha_i = \pm 1/\sqrt{n}$, $\kappa(D) = 1/\sqrt{n}$. $\square$

Una proprietà importante

Limitiamoci a considerare $D = \{\pm e_1, \dots, \pm e_n\}$. Allora

$$\kappa(D) = \min_{v\in\mathbb{R}^n} \max_{d\in D} \frac{v^\top d}{\|v\|\|d\|} = \min_{v\in\mathbb{R}^n} \max_{i=1,\dots,n} \frac{|v^\top e_i|}{\|v\|}$$
$$= \min_{v\in\mathbb{R}^n, \|v\|=1} \max_{i=1,\dots,n} |v^\top e_i|$$

Lemma

$\kappa(D) = 1/\sqrt{n}$

dim. Per ogni $v \in \mathbb{R}^n$, $|v^\top e_i| = |v_i|$. Se $\|v\|^2 = \sum_{i=1}^n |v_i|^2 = 1$, allora $|v^\top e_j| \geq 1/\sqrt{n}$ per qualche j. Quindi, $\kappa(D) \geq 1/\sqrt{n}$.
Se consideriamo $v = \alpha_1 e_1 + \dots + \alpha_n e_n$ con $\alpha_i = \pm 1/\sqrt{n}$,
$\kappa(D) = 1/\sqrt{n}$. □

Una proprietà importante

Limitiamoci a considerare $D = \{\pm e_1, \dots, \pm e_n\}$. Allora

$$\begin{aligned}\kappa(D) &= \min_{v\in\mathbb{R}^n}\max_{d\in D}\frac{v^\top d}{\|v\|\|d\|} = \min_{v\in\mathbb{R}^n}\max_{i=1,\dots,n}\frac{|v^\top e_i|}{\|v\|} \\ &= \min_{v\in\mathbb{R}^n,\|v\|=1}\max_{i=1,\dots,n}|v^\top e_i|\end{aligned}$$

Lemma

$\kappa(D) = 1/\sqrt{n}$

dim. Per ogni $v \in \mathbb{R}^n$, $|v^\top e_i| = |v_i|$. Se $\|v\|^2 = \sum_{i=1}^n |v_i|^2 = 1$, allora $|v^\top e_j| \geq 1/\sqrt{n}$ per qualche j. Quindi, $\kappa(D) \geq 1/\sqrt{n}$.
Se consideriamo $v = \alpha_1 e_1 + \cdots + \alpha_n e_n$ con $\alpha_i = \pm 1/\sqrt{n}$, $\kappa(D) = 1/\sqrt{n}$. $\square$

Una proprietà importante

Limitiamoci a considerare $D = \{\pm e_1, \dots, \pm e_n\}$. Allora

$$\begin{aligned}\kappa(D) &= \min_{v\in\mathbb{R}^n} \max_{d\in D} \frac{v^\top d}{\|v\|\|d\|} = \min_{v\in\mathbb{R}^n} \max_{i=1,\dots,n} \frac{|v^\top e_i|}{\|v\|} \\ &= \min_{v\in\mathbb{R}^n, \|v\|=1} \max_{i=1,\dots,n} |v^\top e_i|\end{aligned}$$

Lemma

$\kappa(D) = 1/\sqrt{n}$

dim. Per ogni $v \in \mathbb{R}^n$, $|v^\top e_i| = |v_i|$. Se $\|v\|^2 = \sum_{i=1}^n |v_i|^2 = 1$, allora $|v^\top e_j| \geq 1/\sqrt{n}$ per qualche j. Quindi, $\kappa(D) \geq 1/\sqrt{n}$. Se consideriamo $v = \alpha_1 e_1 + \dots + \alpha_n e_n$ con $\alpha_i = \pm 1/\sqrt{n}$, $\kappa(D) = 1/\sqrt{n}$. □

Una proprietà importante

Limitiamoci a considerare $D = \{\pm e_1, \dots, \pm e_n\}$. Allora

$$\begin{aligned}\kappa(D) &= \min_{v \in \mathbb{R}^n} \max_{d \in D} \frac{v^\top d}{\|v\| \|d\|} = \min_{v \in \mathbb{R}^n} \max_{i=1,\dots,n} \frac{|v^\top e_i|}{\|v\|} \\ &= \min_{v \in \mathbb{R}^n, \|v\|=1} \max_{i=1,\dots,n} |v^\top e_i|\end{aligned}$$

Lemma

$\kappa(D) = 1/\sqrt{n}$

dim. Per ogni $v \in \mathbb{R}^n$, $|v^\top e_i| = |v_i|$. Se $\|v\|^2 = \sum_{i=1}^n |v_i|^2 = 1$, allora $|v^\top e_j| \geq 1/\sqrt{n}$ per qualche j. Quindi, $\kappa(D) \geq 1/\sqrt{n}$. Se consideriamo $v = \alpha_1 e_1 + \cdots + \alpha_n e_n$ con $\alpha_i = \pm 1/\sqrt{n}$, $\kappa(D) = 1/\sqrt{n}$. $\square$

Convergenza a punti stazionari

Per i metodi visti vale il risultato

Teorema

Se vale (A1) e se $\nabla f(x)$ è Lipschitz continuo modulo L, allora

$$\liminf_{k\to\infty} \|\nabla f(x_k)\| = 0.$$

(Ovvero almeno un punto limite di $\{x_k\}$ è stazionario)

Dim.

- $\{x_k\} \subset L(x_0)$;
- esiste $K_1 \subseteq \{0, 1, 2, \dots\}$ s.t. $\Delta_{k+1} = \Delta_k/2$.

Consideriamo la sottosuccessione $\{x_k\}_{k\in K_1}$ (contenuta in $L(x_0)$, ammette punti limite). Percui, esiste $K_2 \subseteq K_1$ s.t.

- $\lim_{k\to\infty, k\in K_2} x_k = \bar{x}$;
- $\lim_{k\to\infty, k\in K_2} \Delta_k = 0$.

Convergenza a punti stazionari

Per i metodi visti vale il risultato

Teorema

Se vale (A1) e se $\nabla f(x)$ è Lipschitz continuo modulo L, allora

$$\liminf_{k\to\infty} \|\nabla f(x_k)\| = 0.$$

(Ovvero almeno un punto limite di $\{x_k\}$ è stazionario)

Dim.

- $\{x_k\} \subset L(x_0)$;
- esiste $K_1 \subseteq \{0, 1, 2, \dots\}$ s.t. $\Delta_{k+1} = \Delta_k/2$.

Consideriamo la sottosuccessione $\{x_k\}_{k\in K_1}$ (contenuta in $L(x_0)$, ammette punti limite). Percui, esiste $K_2 \subseteq K_1$ s.t.

- $\lim_{k\to\infty, k\in K_2} x_k = \bar{x}$;
- $\lim_{k\to\infty, k\in K_2} \Delta_k = 0$.

Convergenza a punti stazionari

Per i metodi visti vale il risultato

Teorema

Se vale (A1) e se $\nabla f(x)$ è Lipschitz continuo modulo L, allora

$$\liminf_{k\to\infty} \|\nabla f(x_k)\| = 0.$$

(Ovvero almeno un punto limite di $\{x_k\}$ è stazionario)

Dim.

- $\{x_k\} \subset L(x_0)$;
- esiste $K_1 \subseteq \{0, 1, 2, \dots\}$ s.t. $\Delta_{k+1} = \Delta_k/2$.

Consideriamo la sottosuccessione $\{x_k\}_{k\in K_1}$ (contenuta in $L(x_0)$, ammette punti limite). Percui, esiste $K_2 \subseteq K_1$ s.t.

- $\lim_{k\to\infty, k\in K_2} x_k = \bar{x}$;
- $\lim_{k\to\infty, k\in K_2} \Delta_k = 0$.

Convergenza a punti stazionari (segue)

Dal fatto che $\kappa(D) = 1/\sqrt{n}$, abbiamo che, per ogni $k \in K_2$, esiste almeno una direzione $d \in D$ tale che

$$\frac{1}{\sqrt{n}}\|\nabla f(x_k)\| \leq -\nabla f(x_k)^\top d \qquad (1)$$

inoltre, siccome l'iterazione $k \in K_2$ è di fallimento:

$$0 \leq f(x_k + \Delta_k d) - f(x_k) = \nabla f(x_k + \beta_k \Delta_k d)^\top \Delta_k d$$

con $\beta_k \in (0, 1)$. Cioè, sommando $-\Delta_k \nabla f(x_k)^\top d$ ad ambo i membri,

$$-\Delta_k \nabla f(x_k)^\top d \leq \Delta_k \left(\nabla f(x_k + \beta_k \Delta_k d) - \nabla f(x_k)\right)^\top d$$

ovvero, per la (1),

$$\frac{1}{\sqrt{n}}\|\nabla f(x_k)\| \leq \left(\nabla f(x_k + \beta_k \Delta_k d) - \nabla f(x_k)\right)^\top d.$$

Convergenza a punti stazionari (segue)

Dal fatto che $\kappa(D) = 1/\sqrt{n}$, abbiamo che, per ogni $k \in K_2$, esiste almeno una direzione $d \in D$ tale che

$$\frac{1}{\sqrt{n}} \|\nabla f(x_k)\| \leq -\nabla f(x_k)^\top d \tag{1}$$

inoltre, siccome l'iterazione $k \in K_2$ è di fallimento:

$$0 \leq f(x_k + \Delta_k d) - f(x_k) = \nabla f(x_k + \beta_k \Delta_k d)^\top \Delta_k d$$

con $\beta_k \in (0,1)$. Cioè, sommando $-\Delta_k \nabla f(x_k)^\top d$ ad ambo i membri,

$$-\Delta_k \nabla f(x_k)^\top d \leq \Delta_k \left(\nabla f(x_k + \beta_k \Delta_k d) - \nabla f(x_k)\right)^\top d$$

ovvero, per la (1),

$$\frac{1}{\sqrt{n}} \|\nabla f(x_k)\| \leq \left(\nabla f(x_k + \beta_k \Delta_k d) - \nabla f(x_k)\right)^\top d.$$

Convergenza a punti stazionari (segue)

Dal fatto che $\kappa(D) = 1/\sqrt{n}$, abbiamo che, per ogni $k \in K_2$, esiste almeno una direzione $d \in D$ tale che

$$\frac{1}{\sqrt{n}}\|\nabla f(x_k)\| \leq -\nabla f(x_k)^\top d \qquad (1)$$

inoltre, siccome l'iterazione $k \in K_2$ è di fallimento:

$$0 \leq f(x_k + \Delta_k d) - f(x_k) = \nabla f(x_k + \beta_k \Delta_k d)^\top \Delta_k d$$

con $\beta_k \in (0, 1)$. Cioè, sommando $-\Delta_k \nabla f(x_k)^\top d$ ad ambo i membri,

$$-\Delta_k \nabla f(x_k)^\top d \leq \Delta_k \left(\nabla f(x_k + \beta_k \Delta_k d) - \nabla f(x_k)\right)^\top d$$

ovvero, per la (1),

$$\frac{1}{\sqrt{n}}\|\nabla f(x_k)\| \leq \left(\nabla f(x_k + \beta_k \Delta_k d) - \nabla f(x_k)\right)^\top d.$$

Convergenza a punti stazionari (segue)

Dal fatto che $\kappa(D) = 1/\sqrt{n}$, abbiamo che, per ogni $k \in K_2$, esiste almeno una direzione $d \in D$ tale che

$$\frac{1}{\sqrt{n}}\|\nabla f(x_k)\| \leq -\nabla f(x_k)^\top d \qquad (1)$$

inoltre, siccome l'iterazione $k \in K_2$ è di fallimento:

$$0 \leq f(x_k + \Delta_k d) - f(x_k) = \nabla f(x_k + \beta_k \Delta_k d)^\top \Delta_k d$$

con $\beta_k \in (0,1)$. Cioè, sommando $-\Delta_k \nabla f(x_k)^\top d$ ad ambo i membri,

$$-\Delta_k \nabla f(x_k)^\top d \leq \Delta_k \left(\nabla f(x_k + \beta_k \Delta_k d) - \nabla f(x_k)\right)^\top d$$

ovvero, per la (1),

$$\frac{1}{\sqrt{n}}\|\nabla f(x_k)\| \leq \left(\nabla f(x_k + \beta_k \Delta_k d) - \nabla f(x_k)\right)^\top d.$$

Convergenza a punti stazionari (segue)

Dal fatto che $\kappa(D) = 1/\sqrt{n}$, abbiamo che, per ogni $k \in K_2$, esiste almeno una direzione $d \in D$ tale che

$$\frac{1}{\sqrt{n}}\|\nabla f(x_k)\| \leq -\nabla f(x_k)^\top d \tag{1}$$

inoltre, siccome l'iterazione $k \in K_2$ è di fallimento:

$$0 \leq f(x_k + \Delta_k d) - f(x_k) = \nabla f(x_k + \beta_k \Delta_k d)^\top \Delta_k d$$

con $\beta_k \in (0,1)$. Cioè, sommando $-\Delta_k \nabla f(x_k)^\top d$ ad ambo i membri,

$$-\Delta_k \nabla f(x_k)^\top d \leq \Delta_k \left(\nabla f(x_k + \beta_k \Delta_k d) - \nabla f(x_k)\right)^\top d$$

ovvero, per la (1),

$$\frac{1}{\sqrt{n}}\|\nabla f(x_k)\| \leq \left(\nabla f(x_k + \beta_k \Delta_k d) - \nabla f(x_k)\right)^\top d.$$

Convergenza a punti stazionari (segue)

Ora, usando la Lipschitz continuità di ∇f

$$\frac{1}{\sqrt{n}}\|\nabla f(x_k)\| \leq L\|\beta_k \Delta_k d\|\|d\| \leq L\Delta_k,$$

quindi

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

Il risultato segue ricordando che $\lim_{k\to\infty, k\in K_2} \Delta_k = 0$ □

Convergenza a punti stazionari (segue)

Ora, usando la Lipschitz continuità di ∇f

$$\frac{1}{\sqrt{n}}\|\nabla f(x_k)\| \leq L\|\beta_k \Delta_k d\|\|d\| \leq L\Delta_k,$$

quindi

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

Il risultato segue ricordando che $\lim_{k\to\infty, k\in K_2} \Delta_k = 0$ $\square$

Convergenza a punti stazionari (segue)

Ora, usando la Lipschitz continuità di ∇f

$$\frac{1}{\sqrt{n}}\|\nabla f(x_k)\| \leq L\|\beta_k \Delta_k d\|\|d\| \leq L\Delta_k,$$

quindi

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

Il risultato segue ricordando che $\lim_{k\to\infty, k\in K_2} \Delta_k = 0$ $\square$

Pseudo-code di "compass search"

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $D = \{\pm e_i, i = 1, \dots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq \texttt{maxit}$ **and** $\Delta \geq \Delta_{min}$ **do**

  $k \leftarrow k + 1$, $y \leftarrow x$

  Let $\bar{d} \in D$ be s.t. $f(x + \Delta \bar{d}) = \min_{d_i \in D} f(x + \Delta d_i)$

  if $f(x + \Delta \bar{d}) < f(x)$ **then**

    $x \leftarrow x + \Delta \bar{d}$

    endif

  end for

  if $f(y) < f(x)$ **then**

    $x \leftarrow y$

  else

    $\Delta \leftarrow \Delta / 2$

  endif

end while

RETURN: x (miglior punto determinato)

Pseudo-code di "compass search" rivisto o debole

INPUT: x_0, Δ_0, Δ_{min}, maxit, $D = \{\pm e_i, i = 1, \dots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq$maxit **and** $\Delta \geq \Delta_{min}$ **do**

  $k \leftarrow k + 1$, $y \leftarrow x$

  if $\exists\ \bar{d} \in D$ s.t. $f(x + \Delta\bar{d}) < f(x)$ **then**

    $x \leftarrow x + \Delta\bar{d}$

    endif

  end for

  if $f(y) < f(x)$ **then**

    $x \leftarrow y$

  else

    $\Delta \leftarrow \Delta/2$

  endif

end while

RETURN: x (miglior punto determinato)

Pseudo-code di "compass search" rivisto o debole

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $D = \{\pm e_i, i = 1, \ldots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq \mathtt{maxit}$ **and** $\Delta \geq \Delta_{min}$ **do**

  $k \leftarrow k + 1$, $y \leftarrow x$

  if $\exists\ \bar{d} \in D$ s.t. $f(y + \Delta\bar{d}) < f(y)$ **then**

    $y \leftarrow y + \Delta\bar{d}$

    endif

  end for

  if $f(y) < f(x)$ **then**

    $x \leftarrow y$

  else

    $\Delta \leftarrow \Delta/2$

  endif

end while

RETURN: x (miglior punto determinato)

Pseudo-code di "compass search" rivisto o debole

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $D = \{\pm e_i, i = 1, \ldots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq \mathtt{maxit}$ **and** $\Delta \geq \Delta_{min}$ **do**

  $k \leftarrow k + 1$, $y \leftarrow x$

  for each $\bar{d} \in D$

    if $f(y + \Delta\bar{d}) < f(y)$ **then**

      $y \leftarrow y + \Delta\bar{d}$, **break**

    endif

  end for

  if $f(y) < f(x)$ **then**

    $x \leftarrow y$

  else

    $\Delta \leftarrow \Delta/2$

  endif

end while

RETURN: x (miglior punto determinato)

Un nuovo metodo

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $D = \{\pm e_i, i = 1, \ldots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq \mathtt{maxit}$ **and** $\Delta \geq \Delta_{min}$ **do**

  $k \leftarrow k + 1$, $y \leftarrow x$

  for each $\bar{d} \in D$

    if $f(y + \Delta\bar{d}) < f(y)$ **then**

      $y \leftarrow y + \Delta\bar{d}$

    endif

  end for

  if $f(y) < f(x)$ **then**

    $x \leftarrow y$

  else

    $\Delta \leftarrow \Delta/2$

  endif

end while

RETURN: x (miglior punto determinato)

Un “nuovo” metodo

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $D = \{\pm e_i, i = 1, \dots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \le \texttt{maxit}$ **and** $\Delta \ge \Delta_{min}$ **do**

$k \leftarrow k + 1$, $y \leftarrow x$

end while

RETURN: x (miglior punto determinato)

Domande?

Un "nuovo" metodo

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $D = \{\pm e_i, i = 1, \ldots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq$maxit **and** $\Delta \geq \Delta_{min}$ **do**

  $k \leftarrow k + 1$, $y \leftarrow x$

  for each $d \in D$

    if $f(y + \Delta d) < f(y)$ then $y \leftarrow y + \Delta d$

  end for

end while

RETURN: x (miglior punto determinato)

Domande?

Un "nuovo" metodo

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $D = \{\pm e_i, i = 1, \dots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq \mathtt{maxit}$ **and** $\Delta \geq \Delta_{min}$ **do**

$\quad k \leftarrow k + 1$, $y \leftarrow x$

$\quad$ **for each** $d \in D$

$\quad\quad$ **if** $f(y + \Delta d) < f(y)$ **then** $y \leftarrow y + \Delta d$

$\quad$ **end for**

$\quad$ **if** $f(y) < f(x)$ **then**

$\quad\quad x \leftarrow y$

$\quad$ **else** $\Delta \leftarrow \Delta/2$

end while

RETURN: x (miglior punto determinato)

Domande?

Un "nuovo" metodo

INPUT: x_0, Δ_0, Δ_{min}, maxit, $D = \{\pm e_i, i = 1, \ldots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq$maxit **and** $\Delta \geq \Delta_{min}$ **do**

  $k \leftarrow k + 1$, $y \leftarrow x$
  for each $d \in D$
    if $f(y + \Delta d) < f(y)$ **then** $y \leftarrow y + \Delta d$
  end for
  if $f(y) < f(x)$ **then**
    $x \leftarrow y$
  else $\Delta \leftarrow \Delta/2$

end while

RETURN: x (miglior punto determinato)

Domande?

Un "nuovo" metodo

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $D = \{\pm e_i, i = 1, \dots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq \mathtt{maxit}$ **and** $\Delta \geq \Delta_{min}$ **do**

$k \leftarrow k + 1$, $y \leftarrow x$
for each $d \in D$
  if $f(y + \Delta d) < f(y)$ **then** $y \leftarrow y + \Delta d$
end for
if $f(y) < f(x)$ **then**
  $x \leftarrow y$
else $\Delta \leftarrow \Delta/2$

end while

RETURN: x (miglior punto determinato)

Domande?

Un "nuovo" metodo

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $D = \{\pm e_i, i = 1, \dots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq \texttt{maxit}$ **and** $\Delta \geq \Delta_{min}$ **do**

> $k \leftarrow k + 1$, $y \leftarrow x$
> **for each** $d \in D$
>   **if** $f(y + \Delta d) < f(y)$ **then** $y \leftarrow y + \Delta d$
> **end for**
> **if** $f(y) < f(x)$ **then**
> $x \leftarrow y$
> **else** $\Delta \leftarrow \Delta/2$

end while

RETURN: x (miglior punto determinato)

Domande?

Il metodo di Hooke&Jeeves

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $D = \{\pm e_i, i = 1, \dots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq \texttt{maxit}$ **and** $\Delta \geq \Delta_{min}$ **do**

$\quad k \leftarrow k + 1$, $y \leftarrow x$

end while

RETURN: x (miglior punto determinato)

Il metodo di Hooke&Jeeves

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $D = \{\pm e_i, i = 1, \ldots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq \mathtt{maxit}$ **and** $\Delta \geq \Delta_{min}$ **do**

$k \leftarrow k + 1$, $y \leftarrow x$

for each $d \in D$ (*exploratory moves from x*)

if $f(y + \Delta d) < f(y)$ **then** $y \leftarrow y + \Delta d$

end for

end while

RETURN: x (miglior punto determinato)

Il metodo di Hooke&Jeeves

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $D = \{\pm e_i, i = 1, \dots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq \texttt{maxit}$ **and** $\Delta \geq \Delta_{min}$ **do**

  $k \leftarrow k + 1$, $y \leftarrow x$

  for each $d \in D$ (*exploratory moves from* x)

    if $f(y + \Delta d) < f(y)$ **then** $y \leftarrow y + \Delta d$

  end for

  if $f(y) < f(x)$ **then**

  else $\Delta \leftarrow \Delta/2$

end while

RETURN: x (miglior punto determinato)

Il metodo di Hooke&Jeeves

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $D = \{\pm e_i, i = 1, \dots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq$ $\mathtt{maxit}$ **and** $\Delta \geq \Delta_{min}$ **do**

$k \leftarrow k + 1$, $y \leftarrow x$

for each $d \in D$ (<u>*exploratory moves from x*</u>)

if $f(y + \Delta d) < f(y)$ **then** $y \leftarrow y + \Delta d$

end for

if $f(y) < f(x)$ **then** (<u>*pattern move along $y - x$*</u>)

$z \leftarrow y + (y - x)$

else $\Delta \leftarrow \Delta/2$

end while

RETURN: x (miglior punto determinato)

Il metodo di Hooke&Jeeves

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $D = \{\pm e_i, i = 1, \ldots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq \texttt{maxit}$ **and** $\Delta \geq \Delta_{min}$ **do**

> $k \leftarrow k + 1$, $y \leftarrow x$
>
> **for each** $d \in D$ (*exploratory moves from x*)
>
> > **if** $f(y + \Delta d) < f(y)$ **then** $y \leftarrow y + \Delta d$
>
> **end for**
>
> **if** $f(y) < f(x)$ **then** (*pattern move along $y - x$*)
>
> > $z \leftarrow y + (y - x)$
> >
> > for each $d \in D$ (exploratory moves from z)
> >
> > > if $f(z + \Delta d) < f(z)$ then $z \leftarrow z + \Delta d$
> >
> > end for
> >
> > if $f(z) < f(y)$ then $x \leftarrow z$ else $x \leftarrow y$
>
> **else** $\Delta \leftarrow \Delta/2$

end while

RETURN: x (miglior punto determinato)

Il metodo di Hooke&Jeeves

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $D = \{\pm e_i, i = 1, \dots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq \texttt{maxit}$ **and** $\Delta \geq \Delta_{min}$ **do**

  $k \leftarrow k + 1$, $y \leftarrow x$
  for each $d \in D$ (*exploratory moves from x*)
    if $f(y + \Delta d) < f(y)$ **then** $y \leftarrow y + \Delta d$
  end for
  if $f(y) < f(x)$ **then** (*pattern move along $y - x$*)
    $z \leftarrow y + (y - x)$
    for each $d \in D$ (*exploratory moves from z*)
      if $f(z + \Delta d) < f(z)$ **then** $z \leftarrow z + \Delta d$
    end for
    if $f(z) < f(y)$ **then** $x \leftarrow z$ **else** $x \leftarrow y$
  else $\Delta \leftarrow \Delta/2$

end while

RETURN: x (miglior punto determinato)

Il metodo di Hooke&Jeeves

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $D = \{\pm e_i, i = 1, \dots, n\}$

$k \leftarrow 0$, $x \leftarrow x_0$, $\Delta \leftarrow \Delta_0$

while $k \leq \mathtt{maxit}$ **and** $\Delta \geq \Delta_{min}$ **do**

$k \leftarrow k + 1$, $y \leftarrow x$
for each $d \in D$ (*exploratory moves from x*)
  if $f(y + \Delta d) < f(y)$ **then** $y \leftarrow y + \Delta d$
end for
if $f(y) < f(x)$ **then** (*pattern move along $y - x$*)
  $z \leftarrow y + (y - x)$
  for each $d \in D$ (*exploratory moves from z*)
    if $f(z + \Delta d) < f(z)$ **then** $z \leftarrow z + \Delta d$
  end for
  if $f(z) < f(y)$ **then** $x \leftarrow z$ **else** $x \leftarrow y$
else $\Delta \leftarrow \Delta/2$

end while

RETURN: x (miglior punto determinato)

Convergenza di Hooke&Jeeves

Nota bene: Come per il "compass search"

- Se il passo non viene più aggiornato, e.g. $\Delta_{k+1} = \Delta_k$, tutti i punti x_k, da un k in poi, appartengono ad una "griglia";
- Se il punto non cambia, e.g. $x_{k+1} = x_k$,

$$f(x_k + \Delta_k d) \geq f(x_k) \text{ per ogni } d \in D.$$

Quindi, se $L(x_0)$ è compatto e f è cont. differenziabile:

- $\lim_{k\to\infty} \Delta_k = 0$
- almeno un punto limite è stazionario
  - $\nabla f(\bar{x}) = 0$
  - $\liminf_{k\to\infty} \|\nabla f(x_k)\| = 0$

Convergenza di Hooke&Jeeves

Nota bene: Come per il "compass search"

- Se il passo non viene più aggiornato, e.g. $\Delta_{k+1} = \Delta_k$, tutti i punti x_k, da un k in poi, appartengono ad una "griglia";
- Se il punto non cambia, e.g. $x_{k+1} = x_k$,

$$f(x_k + \Delta_k d) \geq f(x_k) \text{ per ogni } d \in D.$$

Quindi, se $L(x_0)$ è compatto e f è cont. differenziabile:

- $\lim_{k\to\infty} \Delta_k = 0$
- almeno un punto limite è stazionario
  - $\nabla f(\bar{x}) = 0$
  - $\liminf_{k\to\infty} \|\nabla f(x_k)\| = 0$

Convergenza di Hooke&Jeeves

Nota bene: Come per il "compass search"

- Se il passo non viene più aggiornato, e.g. $\Delta_{k+1} = \Delta_k$, tutti i punti x_k, da un k in poi, appartengono ad una "griglia";
- Se il punto non cambia, e.g. $x_{k+1} = x_k$,

$$f(x_k + \Delta_k d) \geq f(x_k) \text{ per ogni } d \in D.$$

Quindi, se $L(x_0)$ è compatto e f è cont. differenziabile:

- $\lim_{k\to\infty} \Delta_k = 0$
- almeno un punto limite è stazionario
  - $\nabla f(\bar{x}) = 0$
  - $\liminf_{k\to\infty} \|\nabla f(x_k)\| = 0$

Un metodo generale

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $D = \{\pm e_i, i = 1, \dots, n\}$

$k \leftarrow 0$

while $k \leq \mathtt{maxit}$ **and** $\Delta_k \geq \Delta_{min}$ **do**

  if $\boxed{\exists\ y_k \text{ s.t. } f(y_k) \leq f(x_k) - \gamma\Delta_k}$

    $\Delta_{k+1} \leftarrow \Delta_k$

  elseif $\boxed{\exists\ \bar{d} \in D \text{ s.t. } f(x_k + \Delta_k \bar{d}) \leq f(x_k) - \gamma\Delta_k^2}$

    $y_k \leftarrow x_k + \Delta_k \bar{d}$, $\Delta_{k+1} \leftarrow \Delta_k$

  else

    $y_k \leftarrow x_k$, $\Delta_{k+1} \leftarrow \Delta_k/2$

  endif

  $\boxed{\text{Find } x_{k+1} \text{ s.t. } f(x_{k+1}) \leq f(y_k)}$

  $k \leftarrow k + 1$

end while

RETURN: $\{x_k\}, \{\Delta_k\}$ successioni di punti e passi

Un metodo generale

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $D = \{\pm e_i, i = 1, \dots, n\}$

$k \leftarrow 0$

while $k \leq \mathtt{maxit}$ **and** $\Delta_k \geq \Delta_{min}$ **do**

if $\boxed{\exists\ y_k \text{ s.t. } f(y_k) \leq f(x_k) - \gamma\Delta_k}$

$\Delta_{k+1} \leftarrow \Delta_k$

elseif $\boxed{\exists\ \bar{d} \in D \text{ s.t. } f(x_k + \Delta_k \bar{d}) \leq f(x_k) - \gamma\Delta_k^2}$

$y_k \leftarrow x_k + \Delta_k \bar{d}$, $\Delta_{k+1} \leftarrow \Delta_k$

else

$y_k \leftarrow x_k$, $\Delta_{k+1} \leftarrow \Delta_k/2$

endif

$\boxed{\text{Find } x_{k+1} \text{ s.t. } f(x_{k+1}) \leq f(y_k)}$

$k \leftarrow k+1$

end while

RETURN: $\{x_k\}, \{\Delta_k\}$ successioni di punti e passi

Un metodo generale

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $D = \{\pm e_i, i = 1, \dots, n\}$

$k \leftarrow 0$

while $k \leq \texttt{maxit}$ **and** $\Delta_k \geq \Delta_{min}$ **do**

if $\boxed{\exists\ y_k \text{ s.t. } f(y_k) \leq f(x_k) - \gamma\Delta_k}$

$\Delta_{k+1} \leftarrow \Delta_k$

elseif $\boxed{\exists\ \bar{d} \in D \text{ s.t. } f(x_k + \Delta_k \bar{d}) \leq f(x_k) - \gamma\Delta_k^2}$

$y_k \leftarrow x_k + \Delta_k \bar{d}$, $\Delta_{k+1} \leftarrow \Delta_k$

else

$y_k \leftarrow x_k$, $\Delta_{k+1} \leftarrow \Delta_k/2$

endif

$\boxed{\text{Find } x_{k+1} \text{ s.t. } f(x_{k+1}) \leq f(y_k)}$

$k \leftarrow k+1$

end while

RETURN: $\{x_k\}, \{\Delta_k\}$ successioni di punti e passi

Un metodo generale

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $D = \{\pm e_i, i = 1, \dots, n\}$

$k \leftarrow 0$

while $k \leq \mathtt{maxit}$ **and** $\Delta_k \geq \Delta_{min}$ **do**

if $\boxed{\exists\ y_k \text{ s.t. } f(y_k) \leq f(x_k) - \gamma\Delta_k}$

$\Delta_{k+1} \leftarrow \Delta_k$

elseif $\boxed{\exists\ \bar{d} \in D \text{ s.t. } f(x_k + \Delta_k \bar{d}) \leq f(x_k) - \gamma\Delta_k^2}$

$y_k \leftarrow x_k + \Delta_k \bar{d}$, $\Delta_{k+1} \leftarrow \Delta_k$

else

$y_k \leftarrow x_k$, $\Delta_{k+1} \leftarrow \Delta_k/2$

endif

$\boxed{\text{Find } x_{k+1} \text{ s.t. } f(x_{k+1}) \leq f(y_k)}$

$k \leftarrow k + 1$

end while

RETURN: $\{x_k\}, \{\Delta_k\}$ successioni di punti e passi

Un metodo generale

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $D = \{\pm e_i, i = 1, \ldots, n\}$

$k \leftarrow 0$

while $k \leq \texttt{maxit}$ **and** $\Delta_k \geq \Delta_{min}$ **do**

if $\boxed{\exists\ y_k \text{ s.t. } f(y_k) \leq f(x_k) - \gamma\Delta_k}$

$\Delta_{k+1} \leftarrow \Delta_k$

elseif $\boxed{\exists\ \bar{d} \in D \ \alpha \geq \Delta_k \text{ s.t. } f(x_k + \Delta_k \bar{d}) \leq f(x_k) \ - \gamma\Delta_k^2}$

$y_k \leftarrow x_k + \Delta_k \bar{d}$, $\Delta_{k+1} \leftarrow \Delta_k$

else

$y_k \leftarrow x_k$, $\Delta_{k+1} \leftarrow \Delta_k/2$

endif

$\boxed{\text{Find } x_{k+1} \text{ s.t. } f(x_{k+1}) \leq f(y_k)}$

$k \leftarrow k + 1$

end while

RETURN: $\{x_k\}, \{\Delta_k\}$ successioni di punti e passi

Un metodo generale

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $D = \{\pm e_i, i = 1, \dots, n\}$

$k \leftarrow 0$

while $k \leq \texttt{maxit}$ **and** $\Delta_k \geq \Delta_{min}$ **do**

if $\boxed{\exists\ y_k \text{ s.t. } f(y_k) \leq f(x_k) - \gamma\Delta_k}$

$\Delta_{k+1} \leftarrow \Delta_k$

elseif $\boxed{\exists\ \bar{d} \in D\ \alpha \geq \Delta_k \text{ s.t. } f(x_k + \Delta_k \bar{d}) \leq f(x_k) - \gamma\alpha^2}$

$y_k \leftarrow x_k + \alpha\bar{d}$, $\Delta_{k+1} \leftarrow \alpha$

else

$y_k \leftarrow x_k$, $\Delta_{k+1} \leftarrow \Delta_k/2$

endif

$\boxed{\text{Find } x_{k+1} \text{ s.t. } f(x_{k+1}) \leq f(y_k)}$

$k \leftarrow k + 1$

end while

RETURN: $\{x_k\}, \{\Delta_k\}$ successioni di punti e passi