Ottimizzazione Non Lineare

G. Liuzzi[1]

Mercoledì 22 Novembre 2017

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

Compass Search (forte)

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $D = \{\pm e_1, \cdots \pm e_n\} = \{d_1, \dots, d_m\}$

$k \leftarrow 0$

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

  Let $\bar{d}$, 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

Compass Search (forte)

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $D = \{\pm e_1, \cdots \pm e_n\} = \{d_1, \ldots, d_m\}$

$k \leftarrow 0$

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

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

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

$x_{k+1} \leftarrow y$, $\Delta_{k+1} \leftarrow \Delta_k$

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

Compass Search (forte)

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $D = \{\pm e_1, \cdots \pm e_n\} = \{d_1, \ldots, d_m\}$

$k \leftarrow 0$

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

$\quad y_k \leftarrow x_k + \Delta_k \bar{d}$, d s.t. $f(x_k + \Delta_k \bar{d}) = \min_{d_i \in D} f(x_k + \Delta_k d_i)$

$\quad$ **if** $f(y_k) < f(x_k)$ **then**

$\quad\quad x_{k+1} \leftarrow y_k$, $\Delta_{k+1} \leftarrow \Delta_k$

$\quad$ **else**

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

$\quad$ **endif**

$\quad k \leftarrow k + 1$

end while

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

Compass Search (forte)

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $D = \{\pm e_1, \cdots \pm e_n\} = \{d_1, \ldots, d_m\}$

$k \leftarrow 0$

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

  $y_k{}^1 \leftarrow x_k + \Delta_k d_1$

  for $i = 2, \ldots, m$

    if $f(x_k + \Delta_k d_i) < f(y_k{}^{i-1})$ **then** $y_k{}^i \leftarrow x_k + \Delta_k d_i$

    else $y_k{}^i \leftarrow y_k{}^{i-1}$

  end for

  if $f(y_k{}^m) < f(x_k)$ **then**

    $x_{k+1} \leftarrow y_k{}^m$, $\Delta_{k+1} \leftarrow \Delta_k$

  else

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

  endif

  $k \leftarrow k + 1$

end while

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

Variante (1) di Compass Search

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $D = \{\pm e_1, \cdots \pm e_n\} = \{d_1, \ldots, d_m\}$

$k \leftarrow 0$

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

  $y_k^1 \leftarrow x_k$

  for $i = 1, \ldots, m$

    if $f(y_k^i + \Delta_k d_i) < f({y_k}^i)$ **then** ${y_k}^{i+1} \leftarrow y_k^i + \Delta_k d_i$

    else ${y_k}^{i+1} \leftarrow {y_k}^i$

  end for

  if $f({y_k}^{m+1}) < f(x_k)$ **then**

    $x_{k+1} \leftarrow {y_k}^{m+1}$, $\Delta_{k+1} \leftarrow \Delta_k$

  else

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

  endif

  $k \leftarrow k + 1$

end while

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

Una nota

Se $\bar{D} \subset \mathbb{R}^n$ è una base di $\mathbb{R}^n$, l'insieme di direzioni

$$D = \bar{D} \cup \{-d : d \in \bar{D}\}$$

è tale per cui $\kappa(D) > 0$

Esempio: $\bar{D} = \{e_1, \dots, e_n\}$. $D = \{\pm e_1, \dots, \pm e_n\}$

Variante (2) di Compass Search

INPUT: x_0, Δ_0, Δ_{min}, maxit, $D = \{e_1, \dots e_n\} = \{d_1, \dots, d_n\}$
$k \leftarrow 0$
while $k \leq$ maxit **and** $\Delta_k \geq \Delta_{min}$ **do**
  $y_k^1 \leftarrow x_k$
  for $i = 1, \dots, n$
    if $f(y_k^i + \Delta_k d_i) < f(y_k^i)$ **then** $y_k^{i+1} \leftarrow y_k^i + \Delta_k d_i$
    elseif $f(y_k^i - \Delta_k d_i) < f({y_k}^i)$ **then** ${y_k}^{i+1} \leftarrow y_k^i - \Delta_k d_i$
    else ${y_k}^{i+1} \leftarrow {y_k}^i$
  end for
  if $f({y_k}^{n+1}) < f(x_k)$ **then**
    $x_{k+1} \leftarrow {y_k}^{n+1}$, $\Delta_{k+1} \leftarrow \Delta_k$
  else
    $\Delta_{k+1} \leftarrow \Delta_k/2$, $x_{k+1} \leftarrow x_k$
  endif
  $k \leftarrow k+1$
end while
RETURN: $\{x_k\}, \{{y_k}^i\}, \{\Delta_k\}$ successioni di punti e passi

Variante (2) di Compass Search

INPUT: x_0, $\tilde{\Delta}_0$, Δ_{min}, $\texttt{maxit}$, $D = \{e_1, \dots e_n\} = \{d_1, \dots, d_n\}$
$k \leftarrow 0$
while $k \leq \texttt{maxit}$ **and** $\tilde{\Delta}_k \geq \Delta_{min}$ **do**
  $y_k^1 \leftarrow x_k$
  for $i = 1, \dots, n$
    if $f(y_k^i + \tilde{\Delta}_k d_i) < f(y_k^i)$ **then** $y_k^{i+1} \leftarrow y_k^i + \tilde{\Delta}_k d_i$
    elseif $f(y_k^i - \tilde{\Delta}_k d_i) < f({y_k}^i)$ **then** ${y_k}^{i+1} \leftarrow y_k^i - \tilde{\Delta}_k d_i$
    else ${y_k}^{i+1} \leftarrow {y_k}^i$
  end for
  if $f({y_k}^{n+1}) < f(x_k)$ **then**
    $x_{k+1} \leftarrow {y_k}^{n+1}$, $\tilde{\Delta}_{k+1} \leftarrow \tilde{\Delta}_k$
  else
    $\tilde{\Delta}_{k+1} \leftarrow \tilde{\Delta}_k/2$, $x_{k+1} \leftarrow x_k$
  endif
  $k \leftarrow k + 1$
end while
RETURN: $\{x_k\}, \{{y_k}^i\}, \{\tilde{\Delta}_k\}$ successioni di punti e passi

Variante (2) di Compass Search

INPUT: x_0, $\tilde{\Delta}_0{}^i$, Δ_{min}, $\mathtt{maxit}$, $D = \{e_1, \dots e_n\} = \{d_1, \dots, d_n\}$

$k \leftarrow 0$

while $k \leq \mathtt{maxit}$ **and** $\max_i \tilde{\Delta}_k{}^i \geq \Delta_{min}$ **do**

  $y_k^1 \leftarrow x_k$

  for $i = 1, \dots, n$

    if $f(y_k^i + \tilde{\Delta}_k{}^i d_i) < f(y_k^i)$ **then**

      $y_k^{i+1} \leftarrow y_k^i + \tilde{\Delta}_k{}^i d_i$, $\tilde{\Delta}_{k+1}^i \leftarrow \tilde{\Delta}_k^i$

    elseif $f(y_k^i - \tilde{\Delta}_k{}^i d_i) < f(y_k{}^i)$ **then**

      $y_k{}^{i+1} \leftarrow y_k^i - \tilde{\Delta}_k{}^i d_i$, $\tilde{\Delta}_{k+1}^i \leftarrow \tilde{\Delta}_k^i$

    else $y_k{}^{i+1} \leftarrow y_k{}^i$, $\tilde{\Delta}_{k+1}^i \leftarrow \tilde{\Delta}_k^i/2$

  end for

  $x_{k+1} \leftarrow y_k{}^{n+1}$

  $k \leftarrow k + 1$

end while

RETURN: $\{x_k\}, \{y_k{}^i\}, \{\tilde{\Delta}_k{}^i\}$ successioni di punti e passi

Variante (2) di Compass Search

INPUT: x_0, $\tilde\Delta_0^i$, Δ_{min}, $\texttt{maxit}$, $D = \{e_1, \dots e_n\} = \{d_1, \dots, d_n\}$

$k \leftarrow 0$

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

$y_k^1 \leftarrow x_k$

for $i = 1, \dots, n$

if $f(y_k^i + \tilde\Delta_k^i d_i) < f(y_k^i)$ **then**

$\Delta_k^i \leftarrow \tilde\Delta_k^i$, $p_i \leftarrow d_i$, $\tilde\Delta_{k+1}^i \leftarrow \tilde\Delta_k^i$

elseif $f(y_k^i - \tilde\Delta_k^i d_i) < f(y_k{}^i)$ **then**

$\Delta_k^i \leftarrow \tilde\Delta_k^i$, $p_i \leftarrow -d_i$, $\tilde\Delta_{k+1}^i \leftarrow \tilde\Delta_k^i$

else $\Delta_k^i \leftarrow 0$, $p_i \leftarrow d_i$, $\tilde\Delta_{k+1}^i \leftarrow \tilde\Delta_k^i/2$

$y_k^{i+1} \leftarrow y_k^i + \Delta_k^i p_i$

end for

$x_{k+1} \leftarrow y_k{}^{n+1}$

$k \leftarrow k+1$

end while

RETURN: $\{x_k\}, \{y_k{}^i\}, \{\tilde\Delta_k{}^i\}$ successioni di punti e passi

Variante (3) di Compass Search con espansione

INPUT: x_0, $\tilde{\Delta}_0^i$, Δ_{min}, $\mathtt{maxit}$, $D = \{e_1, \dots e_n\} = \{d_1, \dots, d_n\}$

$k \leftarrow 0$

while $k \leq \mathtt{maxit}$ **and** $\max_i \tilde{\Delta}_k{}^i \geq \Delta_{min}$ **do**

$y_k^1 \leftarrow x_k$

for $i = 1, \dots, n$

if $f(y_k^i + \tilde{\Delta}_k^i d_i) < f(y_k^i)$ **then**

$p_k^i \leftarrow d_i$ e $\boxed{\text{calcola } \Delta_k^i \text{ e } \tilde{\Delta}_{k+1}^i \text{ mediante una espansione lungo } p_k^i}$

elseif $f(y_k^i - \tilde{\Delta}_k^i d_i) < f(y_k{}^i)$ **then**

$p_k^i \leftarrow -d_i$ e $\boxed{\text{calcola } \Delta_k^i \text{ e } \tilde{\Delta}_{k+1}^i \text{ mediante una espansione lungo } p_k^i}$

else $\Delta_k^i \leftarrow 0$, $p_k^i \leftarrow d_i$, $\tilde{\Delta}_{k+1}^i \leftarrow \tilde{\Delta}_k^i/2$

$y_k^{i+1} \leftarrow y_k^i + \Delta_k^i p_k^i$

end for

$x_{k+1} \leftarrow y_k{}^{n+1}$

$k \leftarrow k + 1$

end while

RETURN: $\{x_k\}, \{y_k{}^i\}, \{p_k^i\}, \{\tilde{\Delta}_k{}^i\}$

Espansione lungo p_k^i

Determina il più piccolo intero $j \in \{0, 1, 2, \dots\}$ tale che

$$\begin{array}{rcl} f(y_k^i + 2^j \tilde{\Delta}_k^i d_k^i) & < & f(y_k^i), \\ f(y_k^i + 2^{j+1} \tilde{\Delta}_k^i d_k^i) & \geq & f(y_k^i) \end{array}$$

Poni $\Delta_k^i \leftarrow 2^j \tilde{\Delta}_k^i$ e $\tilde{\Delta}_{k+1}^i \leftarrow \Delta_k^i$

Espansione con suff. decremento

INPUT: x_0, $\tilde{\Delta}_0^i$, Δ_{min}, $\texttt{maxit}$, $D = \{e_1, \dots e_n\} = \{d_1, \dots, d_n\}$

$k \leftarrow 0$

while $k \leq$ maxit **and** $\max_i \tilde{\Delta}_k{}^i \geq \Delta_{min}$ **do**

$y_k^1 \leftarrow x_k$

for $i = 1, \dots, n$

if $f(y_k^i + \tilde{\Delta}_k^i d_i) \leq f(y_k^i) - \gamma(\tilde{\Delta}_k^i)^2$ **then**

$p_k^i \leftarrow d_i$ e $\boxed{\text{calcola } \Delta_k^i \text{ e } \tilde{\Delta}_{k+1}^i \text{ mediante una espansione lungo } p_k^i}$

elseif $f(y_k^i - \tilde{\Delta}_k^i d_i) \leq f(y_k{}^i) - \gamma(\tilde{\Delta}_k^i)^2$ **then**

$p_k^i \leftarrow -d_i$ e $\boxed{\text{calcola } \Delta_k^i \text{ e } \tilde{\Delta}_{k+1}^i \text{ mediante una espansione lungo } p_k^i}$

else $\Delta_k^i \leftarrow 0$, $p_k^i \leftarrow d_i$, $\tilde{\Delta}_{k+1}^i \leftarrow \tilde{\Delta}_k^i/2$

$y_k^{i+1} \leftarrow y_k^i + \Delta_k^i p_k^i$

end for

Scegli x_{k+1} s.t. $f(x_{k+1}) \leq f(y_k{}^{n+1})$ p.es. $x_{k+1} \leftarrow y_k{}^{n+1}$

$k \leftarrow k + 1$

end while

RETURN: $\{x_k\}, \{y_k{}^i\}, \{p_k^i\}, \{\tilde{\Delta}_k{}^i\}$

Espansione lungo d_k^i con suff. decremento

Determina il più piccolo intero $j \in \{0, 1, 2, \dots\}$ tale che

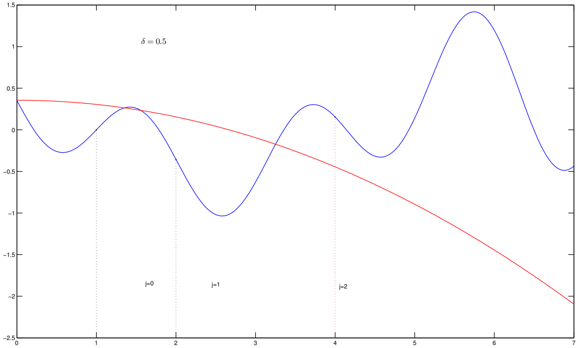
$$\begin{aligned} f(y_k^i + 2^j \tilde{\Delta}_k^i d_k^i) &\leq f(y_k^i) - \gamma(2^j \tilde{\Delta}_k^i)^2, \\ f(y_k^i + 2^{j+1} \tilde{\Delta}_k^i d_k^i) &> f(y_k^i) - \gamma(2^{j+1} \tilde{\Delta}_k^i)^2 \end{aligned}$$

Poni $\Delta_k^i \leftarrow 2^j \tilde{\Delta}_k^i$ e $\tilde{\Delta}_{k+1}^i \leftarrow \Delta_k^i$

Ricerca di linea senza derivate

Siano

$$\begin{aligned}\psi(\beta) &= f(x+\beta d) \text{ e}\\ \Psi(\beta) &= f(x)-\gamma\beta^2\end{aligned}$$



Il metodo è ben posto

Lemma

Se vale la (A1), per ogni iterazione k ed ogni indice $i = 1, \ldots, n$, sono sempre definiti Δ_k^i e $\tilde{\Delta}_{k+1}^i$.

dim. Per dimostrare il lemma è sufficiente far vedere che, quando viene eseguita una espansione lungo d_k^i, esiste sempre un indice $\bar{\jmath}$ finito tale per cui

$$\begin{aligned} f(y_k^i + 2^{\bar{\jmath}}\tilde{\Delta}_k^i d_k^i) &\leq f(y_k^i) - \gamma(2^{\bar{\jmath}}\tilde{\Delta}_k^i)^2, \\ f(y_k^i + 2^{\bar{\jmath}+1}\tilde{\Delta}_k^i d_k^i) &> f(y_k^i) - \gamma(2^{\bar{\jmath}+1}\tilde{\Delta}_k^i)^2. \end{aligned}$$

Supponiamo, per assurdo, che non sia così, e cioè che per ogni $j = 0, 1, 2, \ldots,$

$$f(y_k^i + 2^j\tilde{\Delta}_k^i d_k^i) \leq f(y_k^i) - \gamma(2^j\tilde{\Delta}_k^i)^2.$$

Ma questo è in contraddizione con l'ipotesi (A1). □

Il metodo è ben posto

Lemma

Se vale la (A1), per ogni iterazione k ed ogni indice $i = 1, \dots, n$, sono sempre definiti Δ_k^i e $\tilde{\Delta}_{k+1}^i$.

dim. Per dimostrare il lemma è sufficiente far vedere che, quando viene eseguita una espansione lungo d_k^i, esiste sempre un indice $\bar{\jmath}$ finito tale per cui

$$\begin{aligned} f(y_k^i + 2^{\bar{\jmath}}\tilde{\Delta}_k^i d_k^i) &\leq f(y_k^i) - \gamma(2^{\bar{\jmath}}\tilde{\Delta}_k^i)^2, \\ f(y_k^i + 2^{\bar{\jmath}+1}\tilde{\Delta}_k^i d_k^i) &> f(y_k^i) - \gamma(2^{\bar{\jmath}+1}\tilde{\Delta}_k^i)^2. \end{aligned}$$

Supponiamo, per assurdo, che non sia così, e cioè che per ogni $j = 0, 1, 2, \dots,$

$$f(y_k^i + 2^j\tilde{\Delta}_k^i d_k^i) \leq f(y_k^i) - \gamma(2^j\tilde{\Delta}_k^i)^2.$$

Ma questo è in contraddizione con l'ipotesi (A1). □

Il metodo è ben posto

Lemma

Se vale la (A1), per ogni iterazione k ed ogni indice $i = 1, \dots, n$, sono sempre definiti Δ_k^i e $\tilde{\Delta}_{k+1}^i$.

dim. Per dimostrare il lemma è sufficiente far vedere che, quando viene eseguita una espansione lungo d_k^i, esiste sempre un indice $\bar{\jmath}$ finito tale per cui

$$\begin{aligned} f(y_k^i + 2^{\bar{\jmath}}\tilde{\Delta}_k^i d_k^i) &\leq f(y_k^i) - \gamma(2^{\bar{\jmath}}\tilde{\Delta}_k^i)^2, \\ f(y_k^i + 2^{\bar{\jmath}+1}\tilde{\Delta}_k^i d_k^i) &> f(y_k^i) - \gamma(2^{\bar{\jmath}+1}\tilde{\Delta}_k^i)^2. \end{aligned}$$

Supponiamo, per assurdo, che non sia così, e cioè che per ogni $j = 0, 1, 2, \dots,$

$$f(y_k^i + 2^j\tilde{\Delta}_k^i d_k^i) \leq f(y_k^i) - \gamma(2^j\tilde{\Delta}_k^i)^2.$$

Ma questo è in contraddizione con l'ipotesi (A1). □

Il metodo è ben posto

Lemma

Se vale la (A1), per ogni iterazione k ed ogni indice $i = 1, \dots, n$, sono sempre definiti Δ_k^i e $\tilde{\Delta}_{k+1}^i$.

dim. Per dimostrare il lemma è sufficiente far vedere che, quando viene eseguita una espansione lungo d_k^i, esiste sempre un indice $\bar{\jmath}$ finito tale per cui

$$\begin{array}{rcl} f(y_k^i + 2^{\bar{\jmath}}\tilde{\Delta}_k^i d_k^i) & \leq & f(y_k^i) - \gamma(2^{\bar{\jmath}}\tilde{\Delta}_k^i)^2, \\ f(y_k^i + 2^{\bar{\jmath}+1}\tilde{\Delta}_k^i d_k^i) & > & f(y_k^i) - \gamma(2^{\bar{\jmath}+1}\tilde{\Delta}_k^i)^2. \end{array}$$

Supponiamo, per assurdo, che non sia così, e cioè che per ogni $j = 0, 1, 2, \dots,$

$$f(y_k^i + 2^j\tilde{\Delta}_k^i d_k^i) \leq f(y_k^i) - \gamma(2^j\tilde{\Delta}_k^i)^2.$$

Ma questo è in contraddizione con l'ipotesi (A1). $\square$

Convergenza a zero dei Δ_k^i e $\tilde{\Delta}_k^i$

Proposizione

Se vale la (A1), allora, per ogni $i = 1, \dots, n$, risulta

$$\lim_{k\to\infty} \Delta_k^i = \lim_{k\to\infty} \tilde{\Delta}_k^i = 0.$$

dim. Si vede facilmente che, per ogni k, $f(x_{k+1}) \le f(x)$, quindi $x_k \in L(x_0)$ e $\{f(x_k)\}$ converge.
Risulta inoltre

$$f(x_{k+1}) \le f(x_k) - \gamma \sum_{i=1}^{n} (\Delta_k^i)^2,$$

da cui, prendendo il limite per $k \to \infty$ e ricordando che $\lim_{k\to\infty} f(x_{k+1}) - f(x_k) = 0$,

$$\lim_{k\to\infty} \Delta_k^i = 0, \quad i = 1, \dots, n.$$

Convergenza a zero dei Δ_k^i e $\tilde{\Delta}_k^i$

Proposizione

Se vale la (A1), allora, per ogni $i = 1, \ldots, n$, risulta

$$\lim_{k\to\infty} \Delta_k^i = \lim_{k\to\infty} \tilde{\Delta}_k^i = 0.$$

dim. Si vede facilmente che, per ogni k, $f(x_{k+1}) \leq f(x)$, quindi $x_k \in L(x_0)$ e $\{f(x_k)\}$ converge.

Risulta inoltre

$$f(x_{k+1}) \leq f(x_k) - \gamma \sum_{i=1}^{n} (\Delta_k^i)^2,$$

da cui, prendendo il limite per $k \to \infty$ e ricordando che $\lim_{k\to\infty} f(x_{k+1}) - f(x_k) = 0$,

$$\lim_{k\to\infty} \Delta_k^i = 0, \quad i = 1, \ldots, n.$$

Convergenza a zero dei Δ_k^i e $\tilde{\Delta}_k^i$

Proposizione

Se vale la (A1), allora, per ogni $i = 1, \dots, n$, risulta

$$\lim_{k\to\infty} \Delta_k^i = \lim_{k\to\infty} \tilde{\Delta}_k^i = 0.$$

dim. Si vede facilmente che, per ogni k, $f(x_{k+1}) \le f(x)$, quindi $x_k \in L(x_0)$ e $\{f(x_k)\}$ converge.
Risulta inoltre

$$f(x_{k+1}) \le f(x_k) - \gamma \sum_{i=1}^{n} (\Delta_k^i)^2,$$

da cui, prendendo il limite per $k \to \infty$ e ricordando che $\lim_{k\to\infty} f(x_{k+1}) - f(x_k) = 0$,

$$\lim_{k\to\infty} \Delta_k^i = 0, \quad i = 1, \dots, n.$$

Convergenza a zero dei Δ_k^i e $\tilde{\Delta}_k^i$

Proposizione

Se vale la (A1), allora, per ogni $i = 1, \ldots, n$, risulta

$$\lim_{k\to\infty} \Delta_k^i = \lim_{k\to\infty} \tilde{\Delta}_k^i = 0.$$

dim. Si vede facilmente che, per ogni k, $f(x_{k+1}) \leq f(x)$, quindi $x_k \in L(x_0)$ e $\{f(x_k)\}$ converge.
Risulta inoltre

$$f(x_{k+1}) \leq f(x_k) - \gamma \sum_{i=1}^{n} (\Delta_k^i)^2,$$

da cui, prendendo il limite per $k \to \infty$ e ricordando che $\lim_{k\to\infty} f(x_{k+1}) - f(x_k) = 0$,

$$\lim_{k\to\infty} \Delta_k^i = 0, \quad i = 1, \ldots, n.$$

Convergenza a zero dei Δ_k^i e $\tilde{\Delta}_k^i$ (segue)

Ora, per ogni $i = 1, \dots, n$, è possibile dividere l'insieme $\{0, 1, \dots\}$ in

$$\begin{aligned} K &= \{k : \tilde{\Delta}_{k+1}^i = \Delta_k^i > 0\}, \\ \bar{K} &= \{k : \tilde{\Delta}_{k+1}^i = \tilde{\Delta}_k^i/2, \Delta_k^i = 0\}. \end{aligned}$$

Per ogni $k \in K$ risulta $\Delta_k^i \geq \tilde{\Delta}_k^i$ e quindi

$$\lim_{k\to\infty, k\in K} \tilde{\Delta}_k^i = 0.$$

Per ogni $k \in \bar{K}$, possiamo determinare il più grande indice $m_k < k$ e $m_k \in K$ e scrivere

$$\tilde{\Delta}_k^i = 2^{m_k-k}\Delta_{m_k}^i \leq \Delta_{m_k}^i$$

Convergenza a zero dei Δ_k^i e $\tilde{\Delta}_k^i$ (segue)

Ora, per ogni $i = 1, \dots, n$, è possibile dividere l’insieme $\{0, 1, \dots\}$ in

$$\begin{aligned} K &= \{k : \tilde{\Delta}_{k+1}^i = \Delta_k^i > 0\}, \\ \bar{K} &= \{k : \tilde{\Delta}_{k+1}^i = \tilde{\Delta}_k^i/2, \Delta_k^i = 0\}. \end{aligned}$$

Per ogni $k \in K$ risulta $\Delta_k^i \geq \tilde{\Delta}_k^i$ e quindi

$$\lim_{k \to \infty, k \in K} \tilde{\Delta}_k^i = 0.$$

Per ogni $k \in \bar{K}$, possiamo determinare il più grande indice $m_k < k$ e $m_k \in K$ e scrivere

$$\tilde{\Delta}_k^i = 2^{m_k - k} \Delta_{m_k}^i \leq \Delta_{m_k}^i$$

Convergenza a zero dei Δ_k^i e $\tilde{\Delta}_k^i$ (segue)

Ora, per ogni $i = 1, \dots, n$, è possibile dividere l'insieme $\{0, 1, \dots\}$ in

$$\begin{aligned} K &= \{k : \tilde{\Delta}_{k+1}^i = \Delta_k^i > 0\}, \\ \bar{K} &= \{k : \tilde{\Delta}_{k+1}^i = \tilde{\Delta}_k^i/2, \Delta_k^i = 0\}. \end{aligned}$$

Per ogni $k \in K$ risulta $\Delta_k^i \geq \tilde{\Delta}_k^i$ e quindi

$$\lim_{k\to\infty, k\in K} \tilde{\Delta}_k^i = 0.$$

Per ogni $k \in \bar{K}$, possiamo determinare il più grande indice $m_k < k$ e $m_k \in K$ e scrivere

$$\tilde{\Delta}_k^i = 2^{m_k - k} \Delta_{m_k}^i \leq \Delta_{m_k}^i$$

Convergenza a zero dei Δ_k^i e $\tilde{\Delta}_k^i$ (segue)

$$\tilde{\Delta}_k^i = 2^{m_k - k}\Delta_{m_k}^i \leq \Delta_{m_k}^i$$

Quando $k \to \infty$, $k \in \bar{K}$, possono succedere due cose:

$m_k - k \to -\infty$ (quando K è finito) $\Rightarrow 2^{m_k - k} \to 0$

$m_k \to \infty$ (quando K è infinito) $\Rightarrow \Delta_{m_k}^i \to 0$

Quindi

$$\lim_{k \to \infty, k \in \bar{K}} \tilde{\Delta}_k^i = 0.$$

$\square$

Convergenza a zero dei Δ_k^i e $\tilde{\Delta}_k^i$ (segue)

$$\tilde{\Delta}_k^i = 2^{m_k - k} \Delta_{m_k}^i \leq \Delta_{m_k}^i$$

Quando $k \to \infty$, $k \in \bar{K}$, possono succedere due cose:

$m_k - k \to -\infty$ (quando K è finito) $\Rightarrow 2^{m_k - k} \to 0$

$m_k \to \infty$ (quando K è infinito) $\Rightarrow \Delta_{m_k}^i \to 0$

Quindi

$$\lim_{k \to \infty, k \in \bar{K}} \tilde{\Delta}_k^i = 0.$$

$\square$

Convergenza a zero dei Δ_k^i e $\tilde{\Delta}_k^i$ (segue)

$$\tilde{\Delta}_k^i = 2^{m_k - k} \Delta_{m_k}^i \leq \Delta_{m_k}^i$$

Quando $k \to \infty$, $k \in \bar{K}$, possono succedere due cose:

$m_k - k \to -\infty$ (quando K è finito) $\Rightarrow 2^{m_k - k} \to 0$

$m_k \to \infty$ (quando K è infinito) $\Rightarrow \Delta_{m_k}^i \to 0$

Quindi

$$\lim_{k \to \infty, k \in \bar{K}} \tilde{\Delta}_k^i = 0.$$

$\square$

Convergenza a punti stazionari

Teorema

Se vale la (A1) e se f è continuamente differenziabile, allora

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

dim. Osserviamo prima di tutto che $x_k \in L(x_0)$ quindi $\{x_k\}$ ammette punti limite.
Sia $\bar{x}$ e K tali che

$$\lim_{k\to\infty, k\in K} x_k = \bar{x}.$$

Per ogni $k \in K$ e $i = 2, \dots, n+1$

$$\|y_k^i - x_k\| \le \sum_{j=1}^{i-1} \Delta_k^j \quad \text{quindi} \quad \lim_{k\to\infty, k\in K} y_k^i = \bar{x}.$$

Convergenza a punti stazionari

Teorema

Se vale la (A1) e se f è continuamente differenziabile, allora

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

dim. Osserviamo prima di tutto che $x_k \in L(x_0)$ quindi $\{x_k\}$ ammette punti limite.

Sia $\bar{x}$ e K tali che

$$\lim_{k\to\infty, k\in K} x_k = \bar{x}.$$

Per ogni $k \in K$ e $i = 2, \dots, n+1$

$$\|y_k^i - x_k\| \le \sum_{j=1}^{i-1} \Delta_k^j \quad \text{quindi} \quad \lim_{k\to\infty, k\in K} y_k^i = \bar{x}.$$

Convergenza a punti stazionari

Teorema

Se vale la (A1) e se f è continuamente differenziabile, allora

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

dim. Osserviamo prima di tutto che $x_k \in L(x_0)$ quindi $\{x_k\}$ ammette punti limite.
Sia $\bar{x}$ e K tali che

$$\lim_{k\to\infty, k\in K} x_k = \bar{x}.$$

Per ogni $k \in K$ e $i = 2, \dots, n+1$

$$\|y_k^i - x_k\| \le \sum_{j=1}^{i-1} \Delta_k^j \quad \text{quindi} \quad \lim_{k\to\infty, k\in K} y_k^i = \bar{x}.$$

Convergenza a punti stazionari

Teorema

Se vale la (A1) e se f è continuamente differenziabile, allora

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

dim. Osserviamo prima di tutto che $x_k \in L(x_0)$ quindi $\{x_k\}$ ammette punti limite.
Sia $\bar{x}$ e K tali che

$$\lim_{k\to\infty, k\in K} x_k = \bar{x}.$$

Per ogni $k \in K$ e $i = 2, \dots, n+1$

$$\|y_k^i - x_k\| \le \sum_{j=1}^{i-1} \Delta_k^j \quad \text{quindi} \quad \lim_{k\to\infty, k\in K} y_k^i = \bar{x}.$$

Convergenza a punti stazionari (segue)

Dalla definizione dell'algoritmo segue che uno dei seguenti casi si presenta:

(1) $\Delta^i_k = 0$, $f(y^i_k \pm \tilde\Delta^i_k e_i) > f(y^i_k) - \gamma(\tilde\Delta^i_k)^2$;

(2) $\Delta^i_k > 0$,

$$\begin{array}{cccc} (a) & f(y^i_k + \tilde\Delta^i_k e_i) & > & f(y^i_k) - \gamma(\tilde\Delta^i_k)^2 \\ (b) & f(y^i_k - 2\Delta^i_k e_i) & > & f(y^i_k) - \gamma(2\Delta^i_k)^2; \end{array}$$

(3) $\Delta^i_k > 0$,

$$\begin{array}{cccc} (a) & f(y^i_k + 2\Delta^i_k e_i) & > & f(y^i_k) - \gamma(2\Delta^i_k)^2 \\ (b) & f(y^i_k + \tilde\Delta^i_k e_i) & \le & f(y^i_k) - \gamma(\tilde\Delta^i_k)^2 \end{array}$$

Nel caso (3.b), poniamo $\tilde y^i_k = y^i_k + \tilde\Delta^i_k e_i$, così possiamo scrivere

$$f(\tilde y^i_k - \tilde\Delta^i_k e_i) \ge f(\tilde y^i_k) - \gamma(-(\tilde\Delta^i_k)^2)$$

Convergenza a punti stazionari (segue)

Dalla definizione dell'algoritmo segue che uno dei seguenti casi si presenta:

(1) $\Delta_k^i = 0$, $f(y_k^i \pm \tilde{\Delta}_k^i e_i) > f(y_k^i) - \gamma(\tilde{\Delta}_k^i)^2$;

(2) $\Delta_k^i > 0$,

$$\begin{array}{cccc} (a) & f(y_k^i + \tilde{\Delta}_k^i e_i) & > & f(y_k^i) - \gamma(\tilde{\Delta}_k^i)^2 \\ (b) & f(y_k^i - 2\Delta_k^i e_i) & > & f(y_k^i) - \gamma(2\Delta_k^i)^2; \end{array}$$

(3) $\Delta_k^i > 0$,

$$\begin{array}{cccc} (a) & f(y_k^i + 2\Delta_k^i e_i) & > & f(y_k^i) - \gamma(2\Delta_k^i)^2 \\ (b) & f(y_k^i + \tilde{\Delta}_k^i e_i) & \le & f(y_k^i) - \gamma(\tilde{\Delta}_k^i)^2 \end{array}$$

Nel caso (3.b), poniamo $\tilde{y}_k^i = y_k^i + \tilde{\Delta}_k^i e_i$, così possiamo scrivere

$$f(\tilde{y}_k^i - \tilde{\Delta}_k^i e_i) \ge f(\tilde{y}_k^i) - \gamma(-(\tilde{\Delta}_k^i)^2)$$

Convergenza a punti stazionari (segue)

Dalla definizione dell'algoritmo segue che uno dei seguenti casi si presenta:

(1) $\Delta_k^i = 0$, $f(y_k^i \pm \tilde{\Delta}_k^i e_i) > f(y_k^i) - \gamma(\tilde{\Delta}_k^i)^2$;

(2) $\Delta_k^i > 0$,

$$\begin{array}{lrcl} (a) & f(y_k^i + \tilde{\Delta}_k^i e_i) & > & f(y_k^i) - \gamma(\tilde{\Delta}_k^i)^2 \\ (b) & f(y_k^i - 2\Delta_k^i e_i) & > & f(y_k^i) - \gamma(2\Delta_k^i)^2; \end{array}$$

(3) $\Delta_k^i > 0$,

$$\begin{array}{lrcl} (a) & f(y_k^i + 2\Delta_k^i e_i) & > & f(y_k^i) - \gamma(2\Delta_k^i)^2 \\ (b) & f(y_k^i + \tilde{\Delta}_k^i e_i) & \leq & f(y_k^i) - \gamma(\tilde{\Delta}_k^i)^2 \end{array}$$

Nel caso (3.b), poniamo $\tilde{y}_k^i = y_k^i + \tilde{\Delta}_k^i e_i$, così possiamo scrivere

$$f(\tilde{y}_k^i - \tilde{\Delta}_k^i e_i) \geq f(\tilde{y}_k^i) - \gamma(-(\tilde{\Delta}_k^i)^2)$$

Convergenza a punti stazionari (segue)

Dalla definizione dell'algoritmo segue che uno dei seguenti casi si presenta:

(1) $\Delta_k^i = 0$, $f(y_k^i \pm \tilde{\Delta}_k^i e_i) > f(y_k^i) - \gamma(\tilde{\Delta}_k^i)^2$;

(2) $\Delta_k^i > 0$,

$$\begin{array}{lrcl} (a) & f(y_k^i + \tilde{\Delta}_k^i e_i) & > & f(y_k^i) - \gamma(\tilde{\Delta}_k^i)^2 \\ (b) & f(y_k^i - 2\Delta_k^i e_i) & > & f(y_k^i) - \gamma(2\Delta_k^i)^2; \end{array}$$

(3) $\Delta_k^i > 0$,

$$\begin{array}{lrcl} (a) & f(y_k^i + 2\Delta_k^i e_i) & > & f(y_k^i) - \gamma(2\Delta_k^i)^2 \\ (b) & f(y_k^i + \tilde{\Delta}_k^i e_i) & \leq & f(y_k^i) - \gamma(\tilde{\Delta}_k^i)^2 \end{array}$$

Nel caso (3.b), poniamo $\tilde{y}_k^i = y_k^i + \tilde{\Delta}_k^i e_i$, così possiamo scrivere

$$f(\tilde{y}_k^i - \tilde{\Delta}_k^i e_i) \geq f(\tilde{y}_k^i) - \gamma(-(\tilde{\Delta}_k^i)^2)$$

Convergenza a punti stazionari (segue)

Poniamo

(1) $z_k^i = y_k^i$, $\xi_k^i = \tilde{\Delta}_k^i$;

(2) $z_k^i = y_k^i$,

$$\xi_k^i = \begin{cases} \tilde{\Delta}_k^i & \text{caso (a)} \\ 2\Delta_k^i & \text{caso (b)} \end{cases}$$

(3)

$$z_k^i = \begin{cases} y_k^i & \text{caso (a)} \\ \tilde{y}_k^i & \text{caso (b)} \end{cases} \quad \xi_k^i = \begin{cases} 2\Delta_k^i & \text{caso (a)} \\ \tilde{\Delta}_k^i & \text{caso (b)} \end{cases}$$

E quindi,

$$f(z_k^i \pm \xi_k^i e_i) > f(z_k^i) - \gamma a(\xi_k^i)^2$$

con $z_k^i \to \bar{x}$, $\xi_k^i \to 0$ e $a = 1$ tranne che nel caso (3.b) in cui $a = -1$.

Convergenza a punti stazionari (segue)

Poniamo

(1) $z_k^i = y_k^i$, $\xi_k^i = \tilde{\Delta}_k^i$;

(2) $z_k^i = y_k^i$,

$$\xi_k^i = \begin{cases} \tilde{\Delta}_k^i & \text{caso (a)} \\ 2\Delta_k^i & \text{caso (b)} \end{cases}$$

(3)

$$z_k^i = \begin{cases} y_k^i & \text{caso (a)} \\ \tilde{y}_k^i & \text{caso (b)} \end{cases} \quad \xi_k^i = \begin{cases} 2\Delta_k^i & \text{caso (a)} \\ \tilde{\Delta}_k^i & \text{caso (b)} \end{cases}$$

E quindi,

$$f(z_k^i \pm \xi_k^i e_i) > f(z_k^i) - \gamma a(\xi_k^i)^2$$

con $z_k^i \to \bar{x}$, $\xi_k^i \to 0$ e $a = 1$ tranne che nel caso (3.b) in cui $a = -1$.

Convergenza a punti stazionari (segue)

$$f(z_k^i + \xi_k^i e_i) > f(z_k^i) - \gamma a(\xi_k^i)^2$$
$$f(z_k^i - \xi_k^i e_i) > f(z_k^i) - \gamma a(\xi_k^i)^2$$

Applicando il Teorema della media, risulta

$$\nabla f(u_k^i)^\top e_i > -\gamma a(\xi_k^i)$$
$$\nabla f(v_k^i)^\top (-e_i) > -\gamma a(\xi_k^i)$$

con $u_k^i = z_k^i + t\xi_k^i e_i$, $v_k^i = z_k^i - s\xi_k^i e_i$, $t, s \in (0,1)$.
Prendendo il limite per $k \to \infty, k \in K$, otteniamo

$$\nabla f(\bar{x})^\top e_i \geq 0$$
$$\nabla f(\bar{x})^\top (-e_i) \geq 0$$

da cui segue che $\bar{x}$ è un punto stazionario di f. $\square$

Convergenza a punti stazionari (segue)

$$f(z_k^i + \xi_k^i e_i) > f(z_k^i) - \gamma a(\xi_k^i)^2$$
$$f(z_k^i - \xi_k^i e_i) > f(z_k^i) - \gamma a(\xi_k^i)^2$$

Applicando il Teorema della media, risulta

$$\nabla f(u_k^i)^\top e_i > -\gamma a(\xi_k^i)$$
$$\nabla f(v_k^i)^\top (-e_i) > -\gamma a(\xi_k^i)$$

con $u_k^i = z_k^i + t\xi_k^i e_i,\ v_k^i = z_k^i - s\xi_k^i e_i,\ t, s \in (0,1)$.
Prendendo il limite per $k \to \infty, k \in K$, otteniamo

$$\nabla f(\bar{x})^\top e_i \geq 0$$
$$\nabla f(\bar{x})^\top (-e_i) \geq 0$$

da cui segue che $\bar{x}$ è un punto stazionario di f. $\square$

Convergenza a punti stazionari (segue)

$$f(z_k^i + \xi_k^i e_i) > f(z_k^i) - \gamma a(\xi_k^i)^2$$
$$f(z_k^i - \xi_k^i e_i) > f(z_k^i) - \gamma a(\xi_k^i)^2$$

Applicando il Teorema della media, risulta

$$\nabla f(u_k^i)^\top e_i > -\gamma a(\xi_k^i)$$
$$\nabla f(v_k^i)^\top (-e_i) > -\gamma a(\xi_k^i)$$

con $u_k^i = z_k^i + t\xi_k^i e_i$, $v_k^i = z_k^i - s\xi_k^i e_i$, $t, s \in (0,1)$.
Prendendo il limite per $k \to \infty, k \in K$, otteniamo

$$\nabla f(\bar{x})^\top e_i \geq 0$$
$$\nabla f(\bar{x})^\top (-e_i) \geq 0$$

da cui segue che $\bar{x}$ è un punto stazionario di f. $\square$