

Ottimizzazione Non Lineare

G. Liuzzi¹

Lunedì 20 Novembre 2017

¹Istituto di Analisi dei Sistemi ed Informatica IASI - CNR

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

Un metodo che usa ricerche di linea

INPUT: x_0 , $\tilde{\Delta}_0^i$, $i = 1, \dots, n$, $D = \{e_i, i = 1, \dots, n\}$

for $k = 0, 1, 2, \dots$ **do**

Set $y_k^1 \leftarrow x_k$

for $i = 1, \dots, n$ (*esplora le direzioni in D*)

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

Poni $d_k^i = e_i$ e calcola Δ_k^i e $\tilde{\Delta}_{k+1}^i$ mediante una espansione lungo d_k^i

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

Poni $d_k^i = -e_i$ e calcola Δ_k^i e $\tilde{\Delta}_{k+1}^i$ mediante una espansione lungo d_k^i

else poni $\Delta_k^i \leftarrow 0$, $\tilde{\Delta}_{k+1}^i = \tilde{\Delta}_{k+1}^i/2$

$y_k^{i+1} \leftarrow y_k^i + \Delta_k^i d_k^i$.

end for

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

end for

GENERATE: $\{x_k\}$, $\{\Delta_k^i\}$ and $\{\tilde{\Delta}_k^i\}$, $i = 1, \dots, n$

Espansione lungo d_k^i

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

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

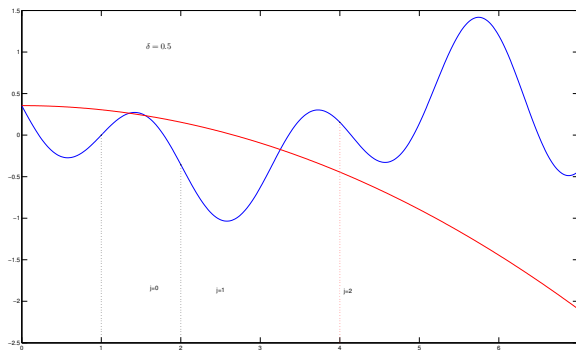
Poni $\Delta_k^i \leftarrow 2^j \tilde{\Delta}_k^i$ e $\tilde{\Delta}_{k+1}^i \leftarrow \Delta_k^i$

Ricerca di linea senza derivate

Siano

$$\psi(\beta) = f(x + \beta d) \text{ e}$$

$$\Psi(\beta) = f(x) - \gamma\beta^2$$



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{j} finito tale per cui

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

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). □

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

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

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 $k \rightarrow \infty$ e ricordando che $\lim_{k \rightarrow \infty} f(x_{k+1}) - f(x_k) = 0$,

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

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

$$K = \{k : \tilde{\Delta}_{k+1}^i = \Delta_k^i > 0\},$$

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

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

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

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

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

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

Quindi

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



Convergenza a punti stazionari

Teorema

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

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

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

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

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

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

Convergenza a punti stazionari (segue)

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

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

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

$$(a) \quad f(y_k^i + \tilde{\Delta}_k^i e_i) > f(y_k^i) - \gamma(\tilde{\Delta}_k^i)^2$$

$$(b) \quad f(y_k^i - 2\Delta_k^i e_i) > f(y_k^i) - \gamma(2\Delta_k^i)^2;$$

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

$$(a) \quad f(y_k^i + 2\Delta_k^i e_i) > f(y_k^i) - \gamma(2\Delta_k^i)^2$$

$$(b) \quad f(y_k^i + \tilde{\Delta}_k^i e_i) \leq f(y_k^i) - \gamma(\tilde{\Delta}_k^i)^2$$

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

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

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

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

(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 $z_k^i \rightarrow \bar{x}$, $\xi_k^i \rightarrow 0$ 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 $k \rightarrow \infty$, $k \in K$, 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 .

