Ottimizzazione dei Sistemi Complessi

G. Liuzzi[1]

Giovedì 3 Marzo 2016

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

Generating Set Search (GSS)

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $0 < \eta_{\max} < 1 < \phi_{\max}$

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

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

  $k \leftarrow k+1$; $D_k = \mathcal{G}_k \cup \mathcal{H}_k$ with

  $PSpan(\mathcal{G}_k) = \mathbb{R}^n$, $|\mathcal{H}_k| < \infty$,
  $\beta_{\min} \leq \|d\| \leq \beta_{\max}$, for all $d \in \mathcal{G}_k$, $\kappa(D_k) > \kappa_{\min}$

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

    $x \leftarrow x + \Delta \bar d$, $\Delta \leftarrow \phi_k \Delta$, $\phi_k \in [1, \phi_{\max}]$

  else

    $\Delta \leftarrow \Delta/2$

  endif

end while

RETURN: x (miglior punto determinato)

Generating Set Search (GSS)

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $0 < \eta_{\max} < 1 < \phi_{\max}$

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

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

$k \leftarrow k+1$; $D_k = \mathcal{G}_k \cup \mathcal{H}_k$ with

$PSpan(\mathcal{G}_k) = \mathbb{R}^n$, $|\mathcal{H}_k| < \infty$,
$\beta_{\min} \leq \|d\| \leq \beta_{\max}$, for all $d \in \mathcal{G}_k$, $\kappa(D_k) > \kappa_{\min}$

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

$x \leftarrow x + \Delta\bar{d}$, $\Delta \leftarrow \phi_k\Delta$, $\phi_k \in [1, \phi_{\max}]$

else

$\Delta \leftarrow \Delta/2$

endif

end while

RETURN: x (miglior punto determinato)

Generating Set Search (GSS)

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $0 < \eta_{\max} < 1 < \phi_{\max}$

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

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

$k \leftarrow k+1$; $D_k = \mathcal{G}_k \cup \mathcal{H}_k$ with

$PSpan(\mathcal{G}_k) = \mathbb{R}^n$, $|\mathcal{H}_k| < \infty$,
$\beta_{\min} \leq \|d\| \leq \beta_{\max}$, for all $d \in \mathcal{G}_k$, $\kappa(D_k) > \kappa_{\min}$

if $\exists\ \bar{d} \in D_k$ s.t. $f(x + \Delta\bar{d}) < f(x)$ **then**
$x \leftarrow x + \Delta\bar{d}$, $\Delta \leftarrow \phi_k\Delta$, $\phi_k \in [1, \phi_{\max}]$
else
$\Delta \leftarrow \Delta/2$
endif

end while

RETURN: x (miglior punto determinato)

Generating Set Search (GSS)

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $0 < \eta_{\max} < 1 < \phi_{\max}$

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

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

$k \leftarrow k+1$; $D_k = \mathcal{G}_k \cup \mathcal{H}_k$ with

$PSpan(\mathcal{G}_k) = \mathbb{R}^n$, $|\mathcal{H}_k| < \infty$,
$\beta_{\min} \leq \|d\| \leq \beta_{\max}$, for all $d \in \mathcal{G}_k$, $\kappa(D_k) > \kappa_{\min}$

if $\exists\ \bar{d} \in D_k$ s.t. $f(x + \Delta\bar{d}) < f(x)$ **then**
$x \leftarrow x + \Delta\bar{d}$, $\Delta \leftarrow \phi_k\Delta$, $\phi_k \in [1, \phi_{\max}]$
else
$\Delta \leftarrow \eta_k\Delta$, $\eta_k \in (0, \eta_{\max}]$
endif

end while

RETURN: x (miglior punto determinato)

Generating Set Search (GSS)

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $0 < \eta_{\max} < 1 < \phi_{\max}$, $\gamma > 0$

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

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

$k \leftarrow k+1$; $D_k = \mathcal{G}_k \cup \mathcal{H}_k$ with

$PSpan(\mathcal{G}_k) = \mathbb{R}^n$, $|\mathcal{H}_k| < \infty$,
$\beta_{\min} \leq \|d\| \leq \beta_{\max}$, for all $d \in \mathcal{G}_k$, $\kappa(D_k) > \kappa_{\min}$

if $\exists\ \bar{d} \in D_k$ s.t. $f(x + \Delta \bar{d}) < f(x)$ **then**
$x \leftarrow x + \Delta \bar{d}$, $\Delta \leftarrow \phi_k \Delta$, $\phi_k \in [1, \phi_{\max}]$
else
$\Delta \leftarrow \eta_k \Delta$, $\eta_k \in (0, \eta_{\max}]$
endif

end while

RETURN: x (miglior punto determinato)

Generating Set Search (GSS)

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $0 < \eta_{\max} < 1 < \phi_{\max}$, $\gamma > 0$

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

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

$k \leftarrow k+1$; $D_k = \mathcal{G}_k \cup \mathcal{H}_k$ with

$PSpan(\mathcal{G}_k) = \mathbb{R}^n$, $|\mathcal{H}_k| < \infty$,
$\beta_{\min} \leq \|d\| \leq \beta_{\max}$, for all $d \in \mathcal{G}_k$, $\kappa(D_k) > \kappa_{\min}$

if $\exists\ \bar{d} \in D_k$ s.t. $f(x + \Delta\bar{d}) \leq f(x) - \gamma\Delta^2$ **then**
$x \leftarrow x + \Delta\bar{d}$, $\Delta \leftarrow \phi_k\Delta$, $\phi_k \geq 1$
else
$\Delta \leftarrow \eta_k\Delta$, $\eta_k \in (0, \eta_{\max}]$
endif

end while

RETURN: x (miglior punto determinato)

Ricerca di linea senza derivate

$$\vdots$$

if $\exists\ d \in D$ t.c. $f(x + \Delta d) \le f(x) - \gamma\Delta^2$ **then**

$\quad x \leftarrow x + \Delta d$

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

$$\vdots$$

Ricerca di linea senza derivate

$$\vdots$$

if $\exists\ d \in D$ t.c. $f(x+\Delta d) \leq f(x) - \gamma\Delta^2$ **then**

$\beta = \min\{\Delta 2^j :\ j = 0, 1, 2, \dots\}$ t.c.

$$\begin{aligned} f(x+\beta d) &\leq f(x) - \gamma\beta^2 \\ f(x+2\beta d) &> f(x) - \gamma(2\beta)^2 \end{aligned}$$

$\Delta \leftarrow \beta,\ x \leftarrow x + \Delta d$

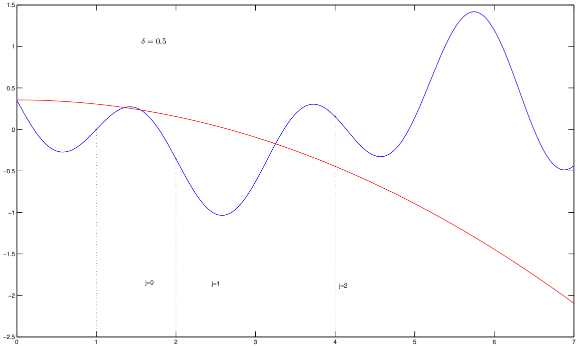
else $\Delta \leftarrow \Delta/2$ **endif**

$$\vdots$$

Ricerca di linea senza derivate

Siano

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



Un metodo che usa ricerche di linea

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $\gamma > 0$,

$D = \{d_1, \dots, d_r\}$ t.c. $PSpan(D) = \mathbb{R}^n$

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

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

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

for $i = 1, 2, \dots, r$

if $f(y + \Delta^i d_i) \leq f(x) - \gamma(\Delta^i)^2$ **then**

else

$\Delta^i \leftarrow \Delta^i/2$

endif

endfor

$x \leftarrow y$

end while

RETURN: x (miglior punto determinato)

Un metodo che usa ricerche di linea

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $\gamma > 0$,
$D = \{d_1, \dots, d_r\}$ t.c. $PSpan(D) = \mathbb{R}^n$

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

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

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

for $i = 1, 2, \dots, r$

if $f(y + \Delta^i d_i) \leq f(x) - \gamma(\Delta^i)^2$ **then**

$y \leftarrow y + \beta d_i$ (β ottenuto con una ric. di linea)

$\Delta^i \leftarrow \beta$

else

$\Delta^i \leftarrow \Delta^i/2$

endif

endfor

$x \leftarrow y$

end while

RETURN: x (miglior punto determinato)

Un metodo che usa ricerche di linea

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $\gamma > 0$,
$D = \{d_1, \dots, d_r\}$ t.c. $PSpan(D) = \mathbb{R}^n$

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

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

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

for $i = 1, 2, \dots, r$

if $f(y + \Delta^i d_i) \leq f(x) - \gamma(\Delta^i)^2$ **then**

$y \leftarrow y + \beta d_i$ (β ottenuto con una ric. di linea)

$\Delta^i \leftarrow \beta$

else

$\Delta^i \leftarrow \Delta^i/2$

endif

endfor

$x \leftarrow y$

end while

RETURN: x (miglior punto determinato)

Un metodo che usa ricerche di linea

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $\gamma > 0$,
$D = \{d_1, \dots, d_r\}$ t.c. $PSpan(D) = \mathbb{R}^n$

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

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

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

  for $i = 1, 2, \dots, r$

    if $f(y + \Delta^i d_i) \leq f(x) - \gamma(\Delta^i)^2$ **then**

      $y \leftarrow y + \beta d_i$ (β ottenuto con una ric. di linea)

      $\Delta^i \leftarrow \beta$

    else

      $\Delta^i \leftarrow \Delta^i/2$

    endif

  endfor

  $x \leftarrow y$

end while

RETURN: x (miglior punto determinato)

Proprietà teoriche

Se $L(x_0)$ è compatto, si può (facilmente) dimostrare che:

- La ricerca di linea restituisce un valore β finito
- $\lim_{k\to\infty} \Delta_k^i = 0$, per ogni $i = 1, \dots, r$
- $\lim_{k\to\infty} \|\nabla f(x_k)\| = 0$ (ogni p.to limite è stazionario)

Proprietà teoriche

Se $L(x_0)$ è compatto, si può (facilmente) dimostrare che:

- La ricerca di linea restituisce un valore β finito
- $\lim_{k\to\infty} \Delta_k^i = 0$, per ogni $i = 1, \ldots, r$
- $\lim_{k\to\infty} \|\nabla f(x_k)\| = 0$ (ogni p.to limite è stazionario)

Proprietà teoriche

Se $L(x_0)$ è compatto, si può (facilmente) dimostrare che:

- La ricerca di linea restituisce un valore β finito
- $\lim_{k\to\infty} \Delta_k^i = 0$, per ogni $i = 1, \ldots, r$
- $\lim_{k\to\infty} \|\nabla f(x_k)\| = 0$ (ogni p.to limite è stazionario)

Proprietà teoriche

Se $L(x_0)$ è compatto, si può (facilmente) dimostrare che:

- La ricerca di linea restituisce un valore β finito
- $\lim_{k\to\infty} \Delta_k^i = 0$, per ogni $i = 1, \dots, r$
- $\lim_{k\to\infty} \|\nabla f(x_k)\| = 0$ (ogni p.to limite è stazionario)

Vi ricordo che ...

... abbiamo considerato il problema:

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

- problema senza vincoli
- $f(x)$ continuamente differenziabile ma
- $\nabla f(x)$ non disponibile (o troppo costoso)
- ricerca di un minimo locale o punto stazionario

Situazioni più complesse

- presenza di vincoli
  - "facili" (limitazioni sulle var.)
  - "difficili" ($g(x) \leq 0$, nonlineari generali)
- $f(x)$ e/o $g(x)$ continue ma non differenziabili
- presenza di più funzioni obiettivo
- presenza di variabili "discrete"
- ricerca di un minimo globale

DFL (Derivative-Free Library), download gratuito di metodi che non usano derivate prime

http://www.iasi.cnr.it/∼liuzzi/DFL

http://www.dis.uniroma1.it/∼lucidi/DFL

Approfondimenti

- Introduction to Derivative-Free Optimization.
  A.R.Conn, K. Scheinberg, L.N. Vicente, MPS-SIAM series on Optimization

Ricostruzione di un codice Julia

Ricostruzione di un codice Julia

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

Ricostruzione di un codice Julia

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