Ottimizzazione Non Lineare

G. Liuzzi[1]

Lunedì 13 Novembre 2017

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

Pseudo-code del metodo "Fermi-Metropolis"

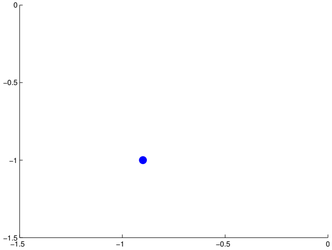
```
INPUT: x_0, Δ_0, Δ_min, maxit
k ← 0, x ← x_0, Δ ← Δ_0
while k ≤ maxit and Δ ≥ Δ_min do
    k ← k + 1, x̃ ← x
    for i = 1, 2, ..., n
        if f(x̃ + Δe_i) < f(x̃) then
            while f(x̃ + Δe_i) < f(x̃) do x̃ ← x̃ + Δe_i end while
        else if f(x̃ − Δe_i) < f(x̃) then
            while f(x̃ − Δe_i) < f(x̃) do x̃ ← x̃ − Δe_i end while
        end if
    end for
    if f(x̃) = f(x) then Δ ← Δ/2
    else x ← x̃
    end if
end while
RETURN: x (miglior punto determinato)
```

Compass Search

Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$

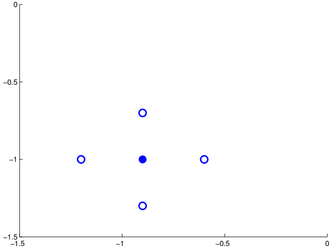
Punto iniziale: $x_0 = (-0.9; -1.0)^\top$
$f(x)$ iniziale: $f(x_0) = 11.3524$
Passo iniziale: $\Delta = 0.3$



Compass Search

Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$



Punto iniziale: $x_0 = (-0.9; -1.0)^\top$
$f(x)$ iniziale: $f(x_0) = 11.3524$
Passo iniziale: $\Delta = 0.3$

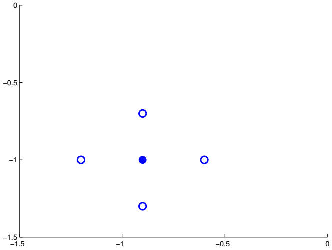
Calcoliamo $f(x)$ nei punti

Est	11.7904
Ovest	19.9504
Nord	
Sud	29.4628

Compass Search

Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$



Punto iniziale: $x_0 = (-0.9; -1.0)^\top$
$f(x)$ iniziale: $f(x_0) = 11.3524$
Passo iniziale: $\Delta = 0.3$

Calcoliamo $f(x)$ nei punti

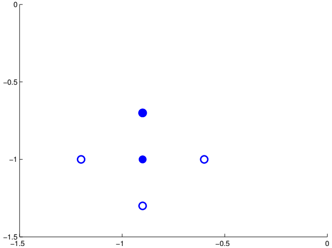
Est	11.7904
Ovest	19.9504
Nord	5.0788
Sud	29.4628

Compass Search

Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$

Punto corrente: $x_k = (-0.9; -0.7)^\top$
$f(x)$ corrente: $f(x_k) = 5.0788$
Passo corrente: $\Delta = 0.3$



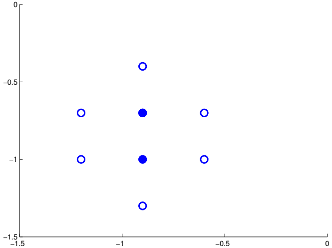
Calcoliamo $f(x)$ nei punti

Est	11.7904
Ovest	19.9504
Nord	**5.0788**
Sud	29.4628

Compass Search

Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$



Punto corrente: $x_k = (-0.9; -0.7)^\top$
$f(x)$ corrente: $f(x_k) = 5.0788$
Passo corrente: $\Delta = 0.3$

Calcoliamo $f(x)$ nei punti

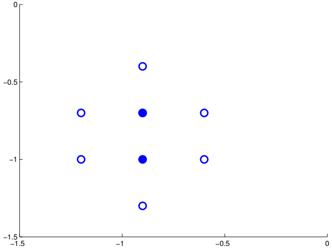
Est	
Ovest	17.4208
Nord	6.4948
Sud	11.3524

Compass Search

Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$

Punto corrente: $x_k = (-0.9; -0.7)^\top$
$f(x)$ corrente: $f(x_k) = 5.0788$
Passo corrente: $\Delta = 0.3$



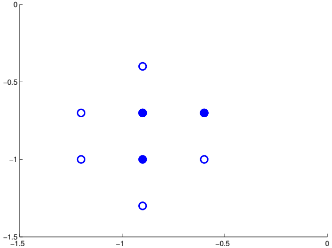
Calcoliamo $f(x)$ nei punti

Est	2.2048
Ovest	17.4208
Nord	6.4948
Sud	11.3524

Compass Search

Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$



Punto corrente: $x_k = (-0.6; -0.7)^\top$
$f(x)$ corrente: $f(x_k) = 2.2048$
Passo corrente: $\Delta = 0.3$

Calcoliamo $f(x)$ nei punti

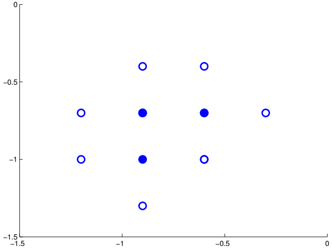
Est	**2.2048**
Ovest	17.4208
Nord	6.4948
Sud	11.3524

Compass Search

Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$

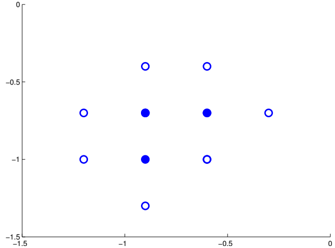
Punto corrente: $x_k = (-0.6; -0.7)^\top$
$f(x)$ corrente: $f(x_k) = 2.2048$
Passo corrente: $\Delta = 0.3$



Calcoliamo $f(x)$ nei punti

Est	4.9108
Ovest	5.0788
Nord	
Sud	11.7904

Compass Search



Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$

Punto corrente: $x_k = (-0.6; -0.7)^\top$
$f(x)$ corrente: $f(x_k) = 2.2048$
Passo corrente: $\Delta = 0.3$

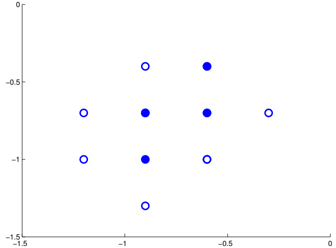
Calcoliamo $f(x)$ nei punti

Est	4.9108
Ovest	5.0788
Nord	0.5248
Sud	11.7904

Compass Search

Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$



Punto corrente: $x_k = (-0.6; -0.4)^\top$
$f(x)$ corrente: $f(x_k) = 0.5248$
Passo corrente: $\Delta = 0.3$

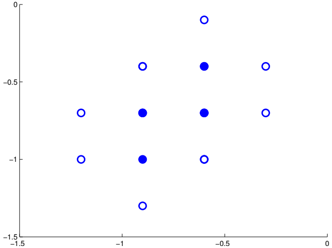
Calcoliamo $f(x)$ nei punti

Est	4.9108
Ovest	5.0788
Nord	**0.5248**
Sud	11.7904

Compass Search

Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$



Punto corrente: $x_k = (-0.6; -0.4)^\top$
$f(x)$ corrente: $f(x_k) = 0.5248$
Passo corrente: $\Delta = 0.3$

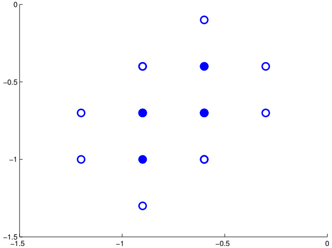
Calcoliamo $f(x)$ nei punti

Est	0.5668
Ovest	6.4948
Nord	3.3808
Sud	2.2048

Compass Search

Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$



Punto corrente: $x_k = (-0.6; -0.4)^\top$
$f(x)$ corrente: $f(x_k) = 0.5248$
Passo corrente: $\Delta = 0.3$

Calcoliamo $f(x)$ nei punti

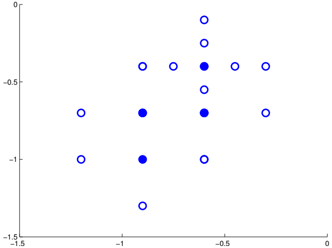
Est	0.5668
Ovest	6.4948
Nord	3.3808
Sud	2.2048

Compass Search

Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$

Punto corrente: $x_k = (-0.6; -0.4)^\top$
$f(x)$ corrente: $f(x_k) = 0.5248$
Passo corrente: $\Delta = 0.15$



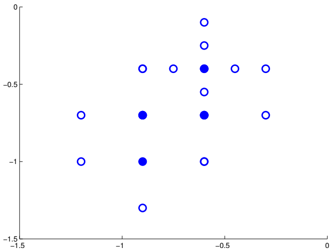
Calcoliamo $f(x)$ nei punti

Est	
Ovest	2.5335
Nord	1.5660
Sud	0.6054

Compass Search

Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$



Punto corrente: $x_k = (-0.6; -0.4)^\top$
$f(x)$ corrente: $f(x_k) = 0.5248$
Passo corrente: $\Delta = 0.15$

Calcoliamo $f(x)$ nei punti

Est	0.0069
Ovest	2.5335
Nord	1.5660
Sud	0.6054

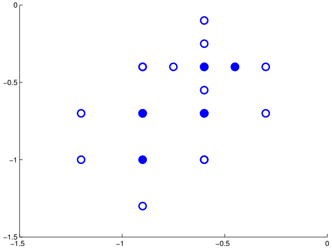
Compass Search

Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$

Punto corrente: $x_k = (-0.45; -0.4)^\top$
$f(x)$ corrente: $f(x_k) = 0.0069$
Passo corrente: $\Delta = 0.15$

Calcoliamo $f(x)$ nei punti

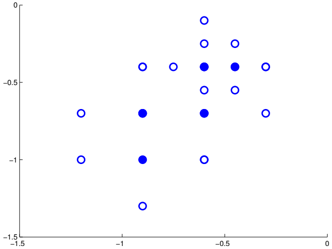


Est	**0.0069**
Ovest	2.5335
Nord	1.5660
Sud	0.6054

Compass Search

Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$



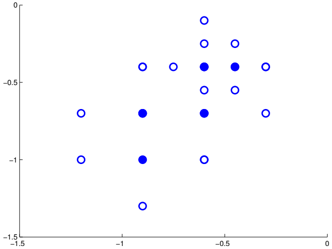
Punto corrente: $x_k = (-0.45; -0.4)^\top$
$f(x)$ corrente: $f(x_k) = 0.0069$
Passo corrente: $\Delta = 0.15$
Calcoliamo $f(x)$ nei punti

Est	0.5668
Ovest	0.5248
Nord	0.3957
Sud	0.7670

Compass Search

Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$



Punto corrente: $x_k = (-0.45; -0.4)^\top$
$f(x)$ corrente: $f(x_k) = 0.0069$
Passo corrente: $\Delta = 0.15$
Calcoliamo $f(x)$ nei punti

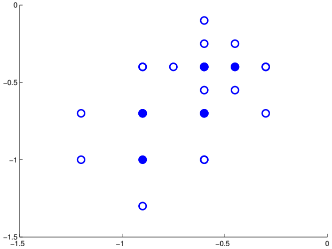
Est	0.5668
Ovest	0.5248
Nord	0.3957
Sud	0.7670

Compass Search

Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$

Punto corrente: $x_k = (-0.45; -0.4)^\top$
$f(x)$ corrente: $f(x_k) = 0.0069$
Passo corrente: $\Delta = 0.075$
Calcoliamo $f(x)$ nei punti

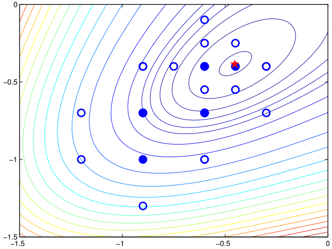


Est	0.5668
Ovest	0.5248
Nord	0.3957
Sud	0.7670

Compass Search

Consideriamo il problema:

$$\min_{x \in \mathbb{R}^2} f(x)$$



Punto corrente: $x_k = (-0.45; -0.4)^\top$
$f(x)$ corrente: $f(x_k) = 0.0069$
Passo corrente: $\Delta = 0.075$
Calcoliamo $f(x)$ nei punti

Est	0.5668
Ovest	0.5248
Nord	0.3957
Sud	0.7670

Ricapitolando

$f(x_k)$	Δ_k
11.352400	0.300000
5.078800	0.300000
2.204800	0.300000
0.524800	0.300000
0.524800	0.150000
0.006925	0.150000
0.006925	0.075000
0.006925	0.037500
0.006925	0.018750
0.000298	0.018750
0.000298	0.009375
0.000298	0.004687
0.000298	0.002344
0.000173	0.002344
0.000054	0.002344
0.000043	0.002344
0.000033	0.002344

Ricapitolando

$f(x_k)$	Δ_k
11.352400	0.300000
5.078800	0.300000
2.204800	0.300000
0.524800	0.300000
0.524800	0.150000
0.006925	0.150000
0.006925	0.075000
0.006925	0.037500
0.006925	0.018750
0.000298	0.018750
0.000298	0.009375
0.000298	0.004687
0.000298	0.002344
0.000173	0.002344
0.000054	0.002344
0.000043	0.002344
0.000033	0.002344

Ricapitolando

$f(x_k)$	Δ_k
11.352400	0.300000
5.078800	0.300000
2.204800	0.300000
0.524800	0.300000
0.524800	0.150000
0.006925	0.150000
0.006925	0.075000
0.006925	0.037500
0.006925	0.018750
0.000298	0.018750
0.000298	0.009375
0.000298	0.004687
0.000298	0.002344
0.000173	0.002344
0.000054	0.002344
0.000043	0.002344
0.000033	0.002344

Ricapitolando

$f(x_k)$	Δ_k
11.352400	0.300000
5.078800	0.300000
2.204800	0.300000
0.524800	0.300000
0.524800	0.150000
0.006925	0.150000
0.006925	0.075000
0.006925	0.037500
0.006925	0.018750
0.000298	0.018750
0.000298	0.009375
0.000298	0.004687
0.000298	0.002344
0.000173	0.002344
0.000054	0.002344
0.000043	0.002344
0.000033	0.002344

Ricapitolando

$f(x_k)$	Δ_k
11.352400	0.300000
5.078800	0.300000
2.204800	0.300000
0.524800	0.300000
0.524800	0.150000
0.006925	0.150000
0.006925	0.075000
0.006925	0.037500
0.006925	0.018750
0.000298	0.018750
0.000298	0.009375
0.000298	0.004687
0.000298	0.002344
0.000173	0.002344
0.000054	0.002344
0.000043	0.002344
0.000033	0.002344

Ricapitolando

$f(x_k)$	Δ_k
11.352400	0.300000
5.078800	0.300000
2.204800	0.300000
0.524800	0.300000
0.524800	0.150000
0.006925	0.150000
0.006925	0.075000
0.006925	0.037500
0.006925	0.018750
0.000298	0.018750
0.000298	0.009375
0.000298	0.004687
0.000298	0.002344
0.000173	0.002344
0.000054	0.002344
0.000043	0.002344
0.000033	0.002344

Ricapitolando

$f(x_k)$	Δ_k
11.352400	0.300000
5.078800	0.300000
2.204800	0.300000
0.524800	0.300000
0.524800	0.150000
0.006925	0.150000
0.006925	0.075000
0.006925	0.037500
0.006925	0.018750
0.000298	0.018750
0.000298	0.009375
0.000298	0.004687
0.000298	0.002344
0.000173	0.002344
0.000054	0.002344
0.000043	0.002344
0.000033	0.002344

Pseudo-code del "compass search"

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $D = \{d_i, i = 1, \ldots, r\}$

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

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

  $k \leftarrow k + 1$

end while

RETURN: x (miglior punto determinato)

Pseudo-code del "compass search"

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $D = \{d_i, i = 1, \ldots, r\}$

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

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

$\quad k \leftarrow k + 1$

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

end while

RETURN: x (miglior punto determinato)

Pseudo-code del "compass search"

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $D = \{d_i, i = 1, \dots, r\}$

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

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

  $k \leftarrow k + 1$

  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**

  else

  endif

end while

RETURN: x (miglior punto determinato)

Pseudo-code del "compass search"

INPUT: x_0, Δ_0, Δ_{min}, $\texttt{maxit}$, $D = \{d_i, i = 1, \dots, r\}$

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

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

  $k \leftarrow k + 1$

  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**

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

  else

    $\Delta \leftarrow \Delta / 2$

  endif

end while

RETURN: x (miglior punto determinato)

Pseudo-code del "compass search"

INPUT: x_0, Δ_0, Δ_{min}, maxit, $D = \{d_i, i = 1, \ldots, r\}$

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

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

$k \leftarrow k + 1$

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**

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

else

$\Delta \leftarrow \Delta/2$

endif

end while

RETURN: x (miglior punto determinato)

Pseudo-code del "compass search"

INPUT: x_0, Δ_0, Δ_{min}, $\mathtt{maxit}$, $D = \{d_i, i = 1, \ldots, r\}$

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

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

  $k \leftarrow k + 1$

  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**

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

  else

    $\Delta \leftarrow \Delta/2$

  endif

end while

RETURN: x (miglior punto determinato)

Pseudo-code del "compass search"

INPUT: x_0, Δ_0, Δ_{min}, maxit, $D = \{d_i, i = 1, \dots, r\}$

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

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

$k \leftarrow k + 1$

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**

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

else

$\Delta \leftarrow \Delta/2$

endif

end while

RETURN: x (miglior punto determinato)

Pseudo-code del "compass search"

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 $k \le \mathtt{maxit}$ **and** $\Delta_k \ge \Delta_{min}$ **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-code del "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** $\Delta_k \geq \Delta_{min}$ **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

Assunzione (A1)

L'insieme di livello $L(x_0) = \{x \in \mathbb{R}^n : \; f(x) \leq f(x_0)\}$ *è compatto*

N.B. nei (due) metodi visti, il passo Δ_k, iterazione dopo iterazione,

- o diminuisce (di un fattore costante $\Delta_{k+1} \leftarrow \Delta_k/2$)
- o rimane costante ($\Delta_{k+1} \leftarrow \Delta_k$)

Lemma

Se (A1) è soddisfatta, $\Delta_{min} \leq 0$ *e* $\texttt{maxit} = +\infty$*, allora* $\{\Delta_k\}$ *è una succ. infinita e*

$$\lim_{k\to\infty} \Delta_k = 0.$$

Un po' di analisi

Assunzione (A1)

L'insieme di livello $L(x_0) = \{x \in \mathbb{R}^n : \; f(x) \leq f(x_0)\}$ *è compatto*

N.B. nei (due) metodi visti, il passo Δ_k, iterazione dopo iterazione,

- o diminuisce (di un fattore costante $\Delta_{k+1} \leftarrow \Delta_k/2$)
- o rimane costante ($\Delta_{k+1} \leftarrow \Delta_k$)

Lemma

Se (A1) è soddisfatta, $\Delta_{min} \leq 0$ *e* $\mathtt{maxit} = +\infty$*, allora* $\{\Delta_k\}$ *è una succ. infinita e*

$$\lim_{k\to\infty} \Delta_k = 0.$$

Un po' di analisi

Assunzione (A1)

L'insieme di livello $L(x_0) = \{x \in \mathbb{R}^n : \ f(x) \leq f(x_0)\}$ *è compatto*

N.B. nei (due) metodi visti, il passo Δ_k, iterazione dopo iterazione,

- o diminuisce (di un fattore costante $\Delta_{k+1} \leftarrow \Delta_k/2$)
- o rimane costante ($\Delta_{k+1} \leftarrow \Delta_k$)

Lemma

Se (A1) è soddisfatta, $\Delta_{min} \leq 0$ *e* $\texttt{maxit} = +\infty$*, allora* $\{\Delta_k\}$ *è una succ. infinita e*

$$\lim_{k\to\infty} \Delta_k = 0.$$

Convergenza a zero del passo

Lemma

Se (A1) è soddisfatta, $\Delta_{min} \leq 0$ e $\texttt{maxit} = +\infty$, allora $\{\Delta_k\}$ è una succ. infinita e

$$\lim_{k\to\infty} \Delta_k = 0.$$

Dim.: Se $\Delta_{min} \leq 0$ e $\texttt{maxit} = +\infty$, la condizione del **while** più esterno è sempre vera, quindi l'algoritmo non termina mai, producendo una succ. $\{\Delta_k\}$ infinita.

Dalla def. dell'algoritmo segue che, per ogni k,

- $\Delta_k > 0$;
- $\Delta_{k+1} \leq \Delta_k$.

Quindi, $\lim_{k\to\infty} \Delta_k = \bar{\Delta} \geq 0$. Supponiamo che $\bar{\Delta} > 0$.

Convergenza a zero del passo

Lemma

Se (A1) è soddisfatta, $\Delta_{min} \leq 0$ e $\texttt{maxit} = +\infty$, allora $\{\Delta_k\}$ è una succ. infinita e

$$\lim_{k\to\infty} \Delta_k = 0.$$

Dim.: Se $\Delta_{min} \leq 0$ e $\texttt{maxit} = +\infty$, la condizione del **while** più esterno è sempre vera, quindi l'algoritmo non termina mai, producendo una succ. $\{\Delta_k\}$ infinita.

Dalla def. dell'algoritmo segue che, per ogni k,

- $\Delta_k > 0$;
- $\Delta_{k+1} \leq \Delta_k$.

Quindi, $\lim_{k\to\infty} \Delta_k = \bar{\Delta} \geq 0$. Supponiamo che $\bar{\Delta} > 0$.

Convergenza a zero del passo

Lemma

Se (A1) è soddisfatta, $\Delta_{min} \leq 0$ *e* $\texttt{maxit} = +\infty$, *allora* $\{\Delta_k\}$ *è una succ. infinita e*

$$\lim_{k\to\infty} \Delta_k = 0.$$

Dim.: Se $\Delta_{min} \leq 0$ e $\texttt{maxit} = +\infty$, la condizione del **while** più esterno è sempre vera, quindi l'algoritmo non termina mai, producendo una succ. $\{\Delta_k\}$ infinita.

Dalla def. dell'algoritmo segue che, per ogni k,

- $\Delta_k > 0$;
- $\Delta_{k+1} \leq \Delta_k$.

Quindi, $\lim_{k\to\infty} \Delta_k = \bar{\Delta} \geq 0$. Supponiamo che $\bar{\Delta} > 0$.

Convergenza a zero del passo (segue)

Esiste un intero $\bar{k}$ tale che, per $k \geq \bar{k}$ il passo non viene più ridotto. Quindi, per ogni $k \geq \bar{k}$

$$f(x_{k+1}) < f(x_k) \text{ e } \Delta_{k+1} = \Delta_k = \Delta_{\bar{k}} = \bar{\Delta}.$$

Quindi, per ogni $k \geq \bar{k}$, i punti x_k appartengono ad una griglia, ovvero, per $i > 0$

$$x_{\bar{k}+i} = x_{\bar{k}} + \bar{\Delta} \sum_{r=0}^{i-1} d_{i_r} \in L(x_0) \qquad d_{i_r} \in D$$

Questo e (A1) implicano che $\{x_k\}_{k \geq \bar{k}}$ (e $\{x_k\}$) è composta da un numero finito di punti distinti.

Convergenza a zero del passo (segue)

Esiste un intero $\bar{k}$ tale che, per $k \geq \bar{k}$ il passo non viene più ridotto. Quindi, per ogni $k \geq \bar{k}$

$$f(x_{k+1}) < f(x_k) \text{ e } \Delta_{k+1} = \Delta_k = \Delta_{\bar{k}} = \bar{\Delta}.$$

Quindi, per ogni $k \geq \bar{k}$, i punti x_k appartengono ad una griglia, ovvero, per $i > 0$

$$x_{\bar{k}+i} = x_{\bar{k}} + \bar{\Delta} \sum_{r=0}^{i-1} d_{i_r} \in L(x_0) \quad d_{i_r} \in D$$

Questo e (A1) implicano che $\{x_k\}_{k \geq \bar{k}}$ (e $\{x_k\}$) è composta da un numero finito di punti distinti.

Convergenza a zero del passo (segue)

Esiste un intero $\bar{k}$ tale che, per $k \geq \bar{k}$ il passo non viene più ridotto. Quindi, per ogni $k \geq \bar{k}$

$$f(x_{k+1}) < f(x_k) \text{ e } \Delta_{k+1} = \Delta_k = \Delta_{\bar{k}} = \bar{\Delta}.$$

Quindi, per ogni $k \geq \bar{k}$, i punti x_k appartengono ad una griglia, ovvero, per $i > 0$

$$x_{\bar{k}+i} = x_{\bar{k}} + \bar{\Delta} \sum_{r=0}^{i-1} d_{i_r} \in L(x_0) \quad d_{i_r} \in D$$

Questo e (A1) implicano che $\{x_k\}_{k \geq \bar{k}}$ (e $\{x_k\}$) è composta da un numero finito di punti distinti.

Convergenza a zero del passo (segue)

Ma allora, deve necessariamente esistere un intero $\tilde{k} > \bar{k}$ t.c.

$$x_{\tilde{k}+1} = x_{\tilde{k}},$$

e quest contraddice il fatto che $f(x_{\tilde{k}+1}) < f(x_{\tilde{k}})$ e allora, $\bar{\Delta} = 0$.

Convergenza a zero del passo (segue)

Ma allora, deve necessariamente esistere un intero $\tilde{k} > \bar{k}$ t.c.

$$x_{\tilde{k}+1} = x_{\tilde{k}},$$

e quest contraddice il fatto che $f(x_{\tilde{k}+1}) < f(x_{\tilde{k}})$ e

allora, $\bar{\Delta} = 0$.

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

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

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

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

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

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.

Convergenza a punti stazionari

Per i metodi visti vale il risultato

Teorema

Se vale (A1) e se $f(x)$ è continuamente differenziabile, 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 $f(x)$ è continuamente differenziabile, 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.

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

Convergenza a punti stazionari

Per i metodi visti vale il risultato

Teorema

Se vale (A1) e se $f(x)$ è continuamente differenziabile, 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)

Per ogni $k \in K_2$ e $i = 1, \dots, n$, abbiamo che

$$\begin{aligned} f(x_k + \Delta_k e_i) &\geq f(x_k) \\ f(x_k - \Delta_k e_i) &\geq f(x_k) \end{aligned}$$

e, per il teorema della media,

$$\begin{aligned} f(x_k + \Delta_k e_i) &= f(x_k) + \Delta_k \nabla f(u_{k,i})^\top e_i \\ f(x_k - \Delta_k e_i) &= f(x_k) - \Delta_k \nabla f(v_{k,i})^\top e_i \end{aligned}$$

dove $u_{k,i} = x_k + \xi_{k,i}\Delta_k e_i$ e $v_{k,i} = x_k - \mu_{k,i}\Delta_k e_i$, con $\xi_{k,i}, \mu_{k,i} \in (0,1)$ e quindi $\lim_{k\to\infty, k\in K_2} u_{k,i} = \lim_{k\to\infty, k\in K_2} v_{k,i} = \bar{x}$.

Convergenza a punti stazionari (segue)

Per ogni $k \in K_2$ e $i = 1, \ldots, n$, abbiamo che

$$\begin{array}{rcl} f(x_k + \Delta_k e_i) & \geq & f(x_k) \\ f(x_k - \Delta_k e_i) & \geq & f(x_k) \end{array}$$

e, per il teorema della media,

$$\begin{array}{rcl} f(x_k + \Delta_k e_i) & = & f(x_k) + \Delta_k \nabla f(u_{k,i})^\top e_i \\ f(x_k - \Delta_k e_i) & = & f(x_k) - \Delta_k \nabla f(v_{k,i})^\top e_i \end{array}$$

dove $u_{k,i} = x_k + \xi_{k,i}\Delta_k e_i$ e $v_{k,i} = x_k - \mu_{k,i}\Delta_k e_i$, con $\xi_{k,i}, \mu_{k,i} \in (0,1)$ e quindi $\displaystyle\lim_{k\to\infty, k\in K_2} u_{k,i} = \lim_{k\to\infty, k\in K_2} v_{k,i} = \bar{x}$.

Convergenza a punti stazionari (segue)

Per ogni $k \in K_2$ e $i = 1, \dots, n$, abbiamo che

$$\begin{array}{rcl} f(x_k + \Delta_k e_i) & \geq & f(x_k) \\ f(x_k - \Delta_k e_i) & \geq & f(x_k) \end{array}$$

e, per il teorema della media,

$$\begin{array}{rcl} f(x_k + \Delta_k e_i) & = & f(x_k) + \Delta_k \nabla f(u_{k,i})^\top e_i \\ f(x_k - \Delta_k e_i) & = & f(x_k) - \Delta_k \nabla f(v_{k,i})^\top e_i \end{array}$$

dove $u_{k,i} = x_k + \xi_{k,i}\Delta_k e_i$ e $v_{k,i} = x_k - \mu_{k,i}\Delta_k e_i$, con $\xi_{k,i}, \mu_{k,i} \in (0,1)$ e quindi $\lim_{k\to\infty, k\in K_2} u_{k,i} = \lim_{k\to\infty, k\in K_2} v_{k,i} = \bar{x}$.

Convergenza a punti stazionari (segue)

Per ogni $k \in K_2$ e $i = 1, \ldots, n$, abbiamo che

$$\begin{array}{rcl} f(x_k + \Delta_k e_i) & \geq & f(x_k) \\ f(x_k - \Delta_k e_i) & \geq & f(x_k) \end{array}$$

e, per il teorema della media,

$$\begin{array}{rcl} f(x_k + \Delta_k e_i) & = & f(x_k) + \Delta_k \nabla f(u_{k,i})^\top e_i \\ f(x_k - \Delta_k e_i) & = & f(x_k) - \Delta_k \nabla f(v_{k,i})^\top e_i \end{array}$$

dove $u_{k,i} = x_k + \xi_{k,i} \Delta_k e_i$ e $v_{k,i} = x_k - \mu_{k,i} \Delta_k e_i$, con $\xi_{k,i}, \mu_{k,i} \in (0,1)$ e quindi $\lim\limits_{k \to \infty, k \in K_2} u_{k,i} = \lim\limits_{k \to \infty, k \in K_2} v_{k,i} = \bar{x}$.

Convergenza a punti stazionari (segue)

Segue anche che:

$$\begin{aligned} \Delta_k \nabla f(u_{k,i})^\top e_i &\geq 0 \\ -\Delta_k \nabla f(v_{k,i})^\top e_i &\geq 0. \end{aligned}$$

Dividendo per Δ_k e prendendo il limite per $k \to \infty, k \in K_2$ otteniamo

$$\begin{aligned} \lim_{k\to\infty, k\in K_2} \nabla f(u_{k,i})^\top e_i &= \nabla f(\bar{x})^\top e_i \geq 0 \\ \lim_{k\to\infty, k\in K_2} -\nabla f(v_{k,i})^\top e_i &= -\nabla f(\bar{x})^\top e_i \geq 0, \end{aligned}$$

e quindi, $\nabla f(\bar{x}) = 0$.

Convergenza a punti stazionari (segue)

Segue anche che:

$$\begin{aligned} \Delta_k \nabla f(u_{k,i})^\top e_i &\geq 0 \\ -\Delta_k \nabla f(v_{k,i})^\top e_i &\geq 0. \end{aligned}$$

Dividendo per Δ_k e prendendo il limite per $k \to \infty, k \in K_2$ otteniamo

$$\begin{aligned} \lim_{k\to\infty, k\in K_2} \nabla f(u_{k,i})^\top e_i &= \nabla f(\bar{x})^\top e_i \geq 0 \\ \lim_{k\to\infty, k\in K_2} -\nabla f(v_{k,i})^\top e_i &= -\nabla f(\bar{x})^\top e_i \geq 0, \end{aligned}$$

e quindi, $\nabla f(\bar{x}) = 0$.

Convergenza a punti stazionari (segue)

Segue anche che:

$$\begin{aligned} \Delta_k \nabla f(u_{k,i})^\top e_i &\geq 0 \\ -\Delta_k \nabla f(v_{k,i})^\top e_i &\geq 0. \end{aligned}$$

Dividendo per Δ_k e prendendo il limite per $k \to \infty, k \in K_2$ otteniamo

$$\begin{aligned} \lim_{k\to\infty, k\in K_2} \nabla f(u_{k,i})^\top e_i &= \nabla f(\bar{x})^\top e_i \geq 0 \\ \lim_{k\to\infty, k\in K_2} -\nabla f(v_{k,i})^\top e_i &= -\nabla f(\bar{x})^\top e_i \geq 0, \end{aligned}$$

e quindi, $\nabla f(\bar{x}) = 0$.

Lipschitz Continuità di ∇f

Definizione

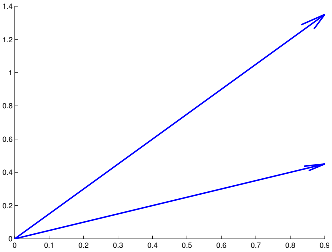
∇f *è Lipschitz continuo (con costante* L*) su un insieme* $A \subseteq \mathbb{R}^n$ *quando, comunque presi due punti* $x, y \in A$*, risulta*

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

Angolo tra due vettori

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

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

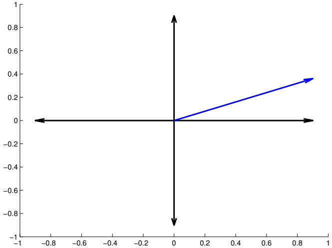


Angolo vettore – insieme di vettori

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

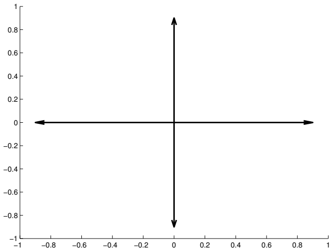
$$\kappa(D) = \min_{v \in \mathbb{R}^n} \phi(v) = \min_{v \in \mathbb{R}^n} \max_{d \in D} \frac{v^\top d}{\|v\|\|d\|}$$

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

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



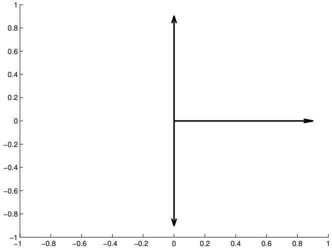
$$\kappa(D) = \frac{1}{\sqrt{2}}$$

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 $D = \{\pm e_1, \dots, \pm e_n\}$. 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 $v = \alpha_1 e_1 + \cdots + \alpha_n e_n$ 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

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

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 $\alpha_i = \pm 1/\sqrt{n}$, $\kappa(D) = 1/\sqrt{n}$. □

Una proprietà importante

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

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

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 $v = \alpha_1 e_1 + \cdots + \alpha_n e_n$ con $\alpha_i = \pm 1/\sqrt{n}$, $\kappa(D) = 1/\sqrt{n}$. □

Una proprietà importante

Limitiamoci a considerare $D = \{\pm e_1, \ldots, \pm e_n\}$. 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 $v = \alpha_1 e_1 + \dots + \alpha_n e_n$ 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

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

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.

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

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

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

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)

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)

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 $\lim_{k\to\infty, k\in K_2} \Delta_k = 0$ $\square$

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 $\lim_{k\to\infty, k\in K_2} \Delta_k = 0$ $\square$