Ottimizzazione Non Lineare

A.A. 2017-18 – 11 Gennaio 2018

prova d'esame (5)

1. Si consideri il seguente insieme in $\mathbb{R}^3$
$$D = \{x \in \mathbb{R}^3 : \ \mathbf{0} \leq x \leq \mathbf{1}\}$$
ed il problema $\min_{x \in D} f(x) = (x_1 - 1/2)^2 + x_1 - 2x_2 - 4x_3$.
   a) Calcolare il valore della funzione obiettivo sul centroide c_0 di D
   b) Applicando la procedura di partizione di `DIRECT` a D, determinare gli insiemi che vengono prodotti.
   c) Tra gli insiemi prodotti al punto precedente, determinare i potenzialmente ottimi quando $f_{min} = -5$ e $\epsilon = 1/5$.

2. Nell'intervallo $[0, 2]$ è definita una funzione incognita $f(x)$ Lipschitz continua di una variabile reale. Sapendo che $L = 5$ è una sovrastima della costante di Lipschitz di f e che $f(0) = 0$ e $f(2) = 4$:
   a) determinare quali valori può assumere $f(x)$ in $x = 1$;
   b) nell'algoritmo di Shubert-Mladineo per la soluzione del problema $\min_{x \in [0,2]} f(x)$, in quale punto si valuta la funzione obiettivo durante la prima iterazione?

3. È dato il seguente problema vincolato nonlineare
$$\begin{array}{lll} \max & e^{x_1} - x_1^2 - x_2^2 & \\ c.v. & x_1^2 + x_2^2 \leq 3 & (\lambda) \\ & 3x_1^3 - 2x_2 = 0 & (\mu) \end{array}$$
   a) Per il problema dato, scrivere l'espressione di una funzione di penalità sequenziale (quadratica) $P(x_1, x_2; \epsilon)$ e del suo gradiente $\nabla_x P(x_1, x_2; \epsilon)$.
   a) Per il problema dato, scrivere l'espressione di una funzione Lagrangiana aumentata sequenziale $L_a(x_1, x_2, \lambda, \mu; \epsilon)$ e del suo gradiente $\nabla L_a(x_1, x_2, \lambda, \mu; \epsilon)$ sia rispetto alle variabili primali x_1, x_2 che rispetto alle variabili duali λ, μ.

4. Si consideri il seguente problema:
$$\begin{array}{ll} \min & e^{x_1} + x_1^2 + x_2^2 \\ c.v. & 0 \leq x_1 \leq 3 \\ & 0 \leq x_2 \leq 3 \end{array}$$
   a) Tenendo conto del fatto che l'insieme ammissibile del problema è convesso, stabilire se il punto $\bar{x} = (0,\ 0)^\top$ è stazionario.
   b) Eseguire una iterazione del metodo di Frank-Wolfe a partire dal punto $\tilde{x} = (2,\ 2)^\top$ ed usando $\gamma = 1/2$, $\delta = 1/2$ nella ricerca di linea di Armijo.

Ottimizzazione Non Lineare

A.A. 2017-18 – 11 Gennaio 2018

prova d'esame (5) – svolgimento

1. a) Il centroide c_0 di D è $c_0 = (1/2,\ 1/2,\ 1/2)^\top$ e $f(c_0) = -5/2$.

   b) Per l'insieme D dato risulta:
$$\delta = 1,\ J = \{1,2,3\},\ m = |J| = 3,$$
quindi vengono generati i $2m = 6$ punti di seguito riportati assieme a c_0
$$\begin{array}{ll}
c_0 = (1/2,\ 1/2,\ 1/2)^\top & f(c_0) = -45/18 \\
c_1 = (5/6,\ 1/2,\ 1/2)^\top & f(c_1) = -37/18 \\
c_2 = (1/6,\ 1/2,\ 1/2)^\top & f(c_2) = -49/18 \\
c_3 = (1/2,\ 5/6,\ 1/2)^\top & f(c_3) = -57/18 \\
c_4 = (1/2,\ 1/6,\ 1/2)^\top & f(c_4) = -33/18 \\
c_5 = (1/2,\ 1/2,\ 5/6)^\top & f(c_5) = \mathbf{-69/18} \\
c_6 = (1/2,\ 1/2,\ 1/6)^\top & f(c_6) = -21/18
\end{array}$$
In grassetto è evidenziato il valore più basso di funzione obiettivo tra i $2m$ valori calcolati. Quindi, la procedura comincia partizionando D per prima lungo la terza coordinata producendo
$$\begin{array}{l}
D^5 = \{x \in \mathbb{R}^3 : 0 \le x_1 \le 1,\ 0 \le x_2 \le 1,\ 2/3 \le x_3 \le 1\}, \\
\bar{D}\ = \{x \in \mathbb{R}^3 : 0 \le x_1 \le 1,\ 0 \le x_2 \le 1,\ 1/3 \le x_3 \le 2/3\}, \\
D^6 = \{x \in \mathbb{R}^3 : 0 \le x_1 \le 1,\ 0 \le x_2 \le 1,\ 0 \le x_3 \le 1/3\}
\end{array}$$
L'insieme $\bar{D}$ contiene, oltre a c_0, i punti
$$\begin{array}{ll}
c_1 = (5/6,\ 1/2,\ 1/2)^\top & f(c_1) = -37/18 \\
c_2 = (1/6,\ 1/2,\ 1/2)^\top & f(c_2) = -49/18 \\
c_3 = (1/2,\ 5/6,\ 1/2)^\top & f(c_3) = \mathbf{-57/18} \\
c_4 = (1/2,\ 1/6,\ 1/2)^\top & f(c_4) = -33/18
\end{array}$$
In grassetto è evidenziato il valore più basso di funzione obiettivo tra i $2(m-1)$ valori residui. La procedura di partizione continua partizionando $\bar{D}$ lungo la coordinata x_2 producendo
$$\begin{array}{l}
D^3 = \{x \in \mathbb{R}^3 : 0 \le x_1 \le 1,\ 2/3 \le x_2 \le 1,\ 1/3 \le x_3 \le 2/3\}, \\
\tilde{D}\ = \{x \in \mathbb{R}^3 : 0 \le x_1 \le 1,\ 1/3 \le x_2 \le 2/3,\ 1/3 \le x_3 \le 2/3\}, \\
D^4 = \{x \in \mathbb{R}^3 : 0 \le x_1 \le 1,\ 0 \le x_2 \le 1/3,\ 1/3 \le x_3 \le 2/3\}
\end{array}$$
L'insieme $\tilde{D}$ contiene ancora, oltre c_0, i due punti c_1 e c_2. A questo punto viene infatti partizionato l'insieme $\tilde{D}$ lungo la coordinata x_1, producendo
$$\begin{array}{l}
D^1 = \{x \in \mathbb{R}^3 : 2/3 \le x_1 \le 1,\ 1/3 \le x_2 \le 2/3,\ 1/3 \le x_3 \le 2/3\}, \\
D^0 = \{x \in \mathbb{R}^3 : 1/3 \le x_1 \le 2/3,\ 1/3 \le x_2 \le 2/3,\ 1/3 \le x_3 \le 2/3\}, \\
D^2 = \{x \in \mathbb{R}^3 : 0 \le x_1 \le 1/3,\ 1/3 \le x_2 \le 2/3,\ 1/3 \le x_3 \le 2/3\}.
\end{array}$$
La partizione ottenuta di D è costituita dagli insiemi $D^0, D^1, D^2, D^3, D^4, D^5$ e D^6.

   c) Indichiamo con $\delta_i = \|u^i - \ell^i\|$ la lunghezza della diagonale maggiore dell'i-esimo insieme D^i. Quindi, abbiamo:
$$\begin{array}{ll}
\delta_0 = \delta_1 = \delta_2 = & \frac{\sqrt{3}}{3} \\
\delta_3 = \delta_4 = & \frac{\sqrt{11}}{3} \\
\delta_5 = \delta_6 = & \frac{\sqrt{19}}{3}.
\end{array}$$
Ora, D^0 e D^1 non possono certamente essere potenzialmente ottimi perché "dominati" da D^2 che ha stessa lunghezza della diagonale ma miglior valore di f.o. nel centroide. Lo stesso vale per D^4, che non può mai essere potenzialmente ottimo perché "dominato" da D^3, e per D^6, che è "dominato" da D^5. Quindi, gli unici insiemi candidati ad essere potenzialmente ottimi sono: D^2, D^3 e D^5.

Consideriamo dapprima D^2. D^2 è potenzialmente ottimo se ammette soluzione $L > 0$ il seguente sistema di disequazioni:

$$\begin{array}{rcl} -\dfrac{49}{18} - L\dfrac{\sqrt{3}}{6} & \leq & -\dfrac{57}{18} - L\dfrac{\sqrt{11}}{6} \\ -\dfrac{49}{18} - L\dfrac{\sqrt{3}}{6} & \leq & -\dfrac{69}{18} - L\dfrac{\sqrt{19}}{6} \\ -\dfrac{49}{18} - L\dfrac{\sqrt{3}}{6} & \leq & -6 \end{array}$$

La prima disequazione implica $L(\sqrt{11} - \sqrt{3})/6 \leq -8/18$ ovvero L minore od uguale ad una quantità negativa. Quindi D^2 non è potenzialmente ottimo.

Consideriamo ora D^3. Affinché D^3 sia potenzialmente ottimo deve ammettere soluzione $L > 0$ il seguente sistema di disequazioni:

$$\begin{array}{rcl} -\dfrac{57}{18} - L\dfrac{\sqrt{11}}{6} & \leq & -\dfrac{49}{18} - L\dfrac{\sqrt{3}}{6} \\ -\dfrac{57}{18} - L\dfrac{\sqrt{11}}{6} & \leq & -\dfrac{69}{18} - L\dfrac{\sqrt{19}}{6} \\ -\dfrac{57}{18} - L\dfrac{\sqrt{11}}{6} & \leq & -6 \end{array}$$

La seconda disequazione implica $L(\sqrt{19} - \sqrt{11})/6 \leq -12/18$ ovvero L minore od uguale ad una quantità negativa. Quindi D^3 non è potenzialmente ottimo.

Pertanto, l'unico insieme potenzialmente ottimo tra i 7 della partizione è D^5.

2. a) Sappiamo che $f(x)$ è Lipschitz continua in $[0, 2]$, che una sovrastima della costante di Lipschitz è $L = 5$ e che $f(0) = 0, f(2) = 4$.

Richiamando la definizione di Lipschitz continuità, sappiamo che, per ogni $x \in [0, 2]$:

$$\begin{array}{rcl} |f(0) - f(x)| & \leq & L|0 - x| \\ |f(2) - f(x)| & \leq & L|2 - x| \end{array}$$

ovvero, per ogni $x \in [0, 2]$

$$\begin{array}{rcccl} -Lx & \leq & f(0) - f(x) & \leq & Lx \\ -L(2 - x) & \leq & f(2) - f(x) & \leq & L(2 - x) \end{array}$$

Quindi,

$$\begin{array}{rcccl} -Lx & \leq & f(x) & \leq & Lx \\ 4 - L(2 - x) & \leq & f(x) & \leq & 4 + L(2 - x) \end{array}$$

cioè, comunque scelto $x \in [0, 2]$, deve risultare

$$\max\{-Lx, 4 - L(2 - x)\} \leq f(x) \leq \min\{Lx, 4 + L(2 - x)\}$$

Se scegliamo $x = 1$ e ricordando che $L = 5$, le disuguaglianze di sopra impongono:

$$-1 \leq f(1) \leq 5$$

b) Nella prima iterazione, l'algoritmo di Shubert-Mladineo determina il punto in cui valutare la funzione obiettivo calcolando il punto di minimo della funzione (lineare a tratti)

$$\ell(x) = \max\{-5x, -6 + 5x\}$$

che è $x = 3/5$.

3. Il problema è in forma di max quindi, di seguito riportiamo il problema in forma di min:

$$\begin{array}{lll} \min & -e^{x_1} + x_1^2 + x_2^2 & \\ c.v. & x_1^2 + x_2^2 \leq 3 & (\lambda) \\ & 3x_1^3 - 2x_2 = 0 & (\mu) \end{array}$$

a)

$$
\begin{aligned}
P(x_1, x_2; \epsilon) &= -e^{x_1} + x_1^2 + x_2^2 + \frac{1}{\epsilon}\Big(\max\{x_1^2 + x_2^2 - 3, 0\}^2 + (3x_1^3 - 2x_2)^2\Big) \\
\nabla P(x_1, x_2; \epsilon) &= \begin{bmatrix} -e^{x_1} + 2x_1 \\ 2x_2 \end{bmatrix} + \frac{2}{\epsilon}\max\{x_1^2 + x_2^2 - 3, 0\}\begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} + \frac{2}{\epsilon}(3x_1^3 - 2x_2)\begin{bmatrix} 9x_1^2 \\ -2 \end{bmatrix}
\end{aligned}
$$

b)

$$
\begin{aligned}
L_a(x_1, x_2, \lambda, \mu; \epsilon) &= -e^{x_1} + x_1^2 + x_2^2 + \mu(3x_1^3 - 2x_2) + \frac{1}{\epsilon}(3x_1^3 - 2x_2)^2 \\
&+ \lambda \max\left\{x_1^2 + x_2^2 - 3, -\frac{\epsilon}{2}\lambda\right\} + \frac{1}{\epsilon}\max\left\{x_1^2 + x_2^2 - 3, -\frac{\epsilon}{2}\lambda\right\}^2 \\
\nabla_x L_a(x_1, x_2, \lambda, \mu; \epsilon) &= \begin{bmatrix} -e^{x_1} + 2x_1 \\ 2x_2 \end{bmatrix} + \left(\mu + \frac{2}{\epsilon}(3x_1^3 - 2x_2)\right)\begin{bmatrix} 9x_1^2 \\ -2 \end{bmatrix} \\
&+ \left(\lambda + \frac{2}{\epsilon}\max\left\{x_1^2 + x_2^2 - 3, -\frac{\epsilon}{2}\lambda\right\}\right)\begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} \\
\nabla_\lambda L_a(x_1, x_2, \lambda, \mu; \epsilon) &= \max\left\{x_1^2 + x_2^2 - 3, -\frac{\epsilon}{2}\lambda\right\} \\
\nabla_\mu L_a(x_1, x_2, \lambda, \mu; \epsilon) &= (3x_1^3 - 2x_2)
\end{aligned}
$$

4. a) Nel punto $\bar{x} = (0,\ 0)^\top$ risulta

$$\nabla f(\bar{x}) = (1, 0)^\top$$

$\bar{x}$ è stazionario quando

$$\nabla f(\bar{x})^\top (x - \bar{x}) \geq 0, \quad \text{per ogni } x \in S$$

cioè quando

$$(1,\ 0)^\top x = x_1 \geq 0, \quad \text{per ogni } x \in S.$$

Siccome ogni x ammissibile ha $x_1 \geq 0$, il punto $\bar{x}$ è stazionario.

b) In $\tilde{x}$ risulta $\nabla f(\tilde{x}) = (e^2 + 4, 4)^\top$. Per eseguire una iterazione del metodo di Frank-Wolfe, come prima cosa bisogna determinare una direzione di ricerca. Per fare questo, risolviamo il sottoproblema (lineare)

$$
\begin{aligned}
&\min\ (e^2 + 4)x_1 + 4x_2 - 2(e^2 + 4) - 8 \\
&c.v.\ 0 \leq x_1 \leq 3 \\
&\qquad 0 \leq x_2 \leq 3
\end{aligned}
$$

La soluzione del problema di sopra è il punto $\hat{x} = (0,\ 0)^\top$ che ci fornisce la direzione di ricerca $d = \hat{x} - \tilde{x} = (-2,\ -2)^\top$.

Sia ora $x(\alpha) = \tilde{x} + \alpha d$, dobbiamo determinare il più piccolo intero non negativo j tale che

$$f(x(\delta^j)) \leq f(\tilde{x}) + \gamma\delta^j \nabla f(\tilde{x})^\top d = e^2 + 8 - \gamma\delta^j(2e^2 + 16).$$

- $\underline{j = 0}$, cioè $\alpha = 1$.

  $$x(1) = (0,\ 0)^\top,\ f(x(1)) = 1 \not\leq 0$$

  quindi il passo $\alpha = 1$ non è accettabile.
- $\underline{j = 1}$, cioè $\alpha = \delta = 1/2$.

  $$x(1/2) = (1,\ 1)^\top,\ f(x(1/2)) = e + 2 \leq (e^2 + 8)/2$$

  Quindi, la ricerca di Armijo restituisce il passo $\alpha = 1/2$ ed il metodo genera il nuovo punto $x_{new} = (1,\ 1)^\top$