Ottimizzazione dei Sistemi Complessi

G. Liuzzi[1]

Mercoledì 25 Marzo 2020

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

Punti di KKT

Definizione (Punto di KKT)

Un punto $x \in \mathbb{R}^n$ *è un* **punto di KKT** *quando esiste un vettore di moltiplicatori (di KKT),* (λ, μ)*, tale che sia soddisfatto il sistema*

$$\begin{aligned} &\nabla f(x) + \nabla g(x)\lambda + \nabla h(x)\mu = \mathbf{0} \\ &\lambda \geq \mathbf{0} \\ &\lambda_i g_i(x) = 0, \quad i = 1, \dots, m \\ &g(x) \leq \mathbf{0}, \ \ h(x) = \mathbf{0} \end{aligned}$$

Problemi convessi

Supponiamo:

- f, g continuamente differenziabili e **convesse**
- h **affini**, e.g. $h_j(x) = a^\top x - b = 0$

N.B. sotto queste ipotesi diciamo che il problema è convesso

$$\begin{array}{l} \min \; x_1 + x_2 \\ s.t. \; x_1^2 + x_2^2 \leq 2 \end{array}$$

non *è convesso*

Problemi convessi

Supponiamo:

- f, g continuamente differenziabili e **convesse**
- h **affini**, e.g. $h_j(x) = a^\top x - b = 0$

N.B. sotto queste ipotesi diciamo che il problema è convesso

$$\begin{array}{l} \min\ x_1 + x_2 \\ s.t.\ x_1^2 + x_2^2 = 2 \end{array}$$

non è convesso

Problemi convessi

Teorema (Ogni minimo locale è globale)

Sia $x^* \in \mathcal{F}$ un punto di **minimo locale** *del problema. allora x^* è anche un* **minimo globale** *del problema.*

Dim. Sia $\mathcal{B}(x^*;\epsilon)$ l'intorno di x^* tale che

$$f(x^*) \leq f(x), \quad \forall\, x \in \mathcal{F} \cap \mathcal{B}(x^*;\epsilon)$$

Supponiamo, per assurdo, che x^* non sia minimo globale, cioè, esiste $x^\circ \in \mathcal{F}$ tale che

$$f(x^\circ) < f(x^*)$$

Problemi convessi

Teorema (Ogni minimo locale è globale)

Sia $x^* \in \mathcal{F}$ un punto di **minimo locale** *del problema. allora x^* è anche un* **minimo globale** *del problema.*

Dim. Sia $\mathcal{B}(x^*;\epsilon)$ l'intorno di x^* tale che

$$f(x^*) \leq f(x), \quad \forall\, x \in \mathcal{F} \cap \mathcal{B}(x^*;\epsilon)$$

Supponiamo, per assurdo, che x^* non sia minimo globale, cioè, esiste $x^\circ \in \mathcal{F}$ tale che

$$f(x^\circ) < f(x^*)$$

Problemi convessi

Teorema (Ogni minimo locale è globale)

Sia $x^* \in \mathcal{F}$ un punto di **minimo locale** *del problema. allora x^* è anche un* **minimo globale** *del problema.*

Dim. Sia $\mathcal{B}(x^*; \epsilon)$ l'intorno di x^* tale che

$$f(x^*) \leq f(x), \quad \forall\, x \in \mathcal{F} \cap \mathcal{B}(x^*; \epsilon)$$

Supponiamo, per assurdo, che x^* non sia minimo globale, cioè, esiste $x^\circ \in \mathcal{F}$ tale che

$$f(x^\circ) < f(x^*)$$

Dimostrazione – continua

Consideriamo il segmento $[x^*, x^\circ]$ che congiunge x^* con x°. Per convessità, risulta $[x^*, x^\circ] \subset \mathcal{F}$ e

$$f(x(\beta)) \le \beta f(x^\circ) + (1-\beta) f(x^*), \quad \forall\, \beta \in [0,1]$$

Deve esistere un valore $\bar\beta \in (0,1)$ tale che il punto

$$x(\bar\beta) = \bar\beta x^\circ + (1-\bar\beta) x^* \in [x^*, x^\circ]$$

appartiene a $\mathcal{F} \cap \mathcal{B}(x^*; \epsilon)$ e quindi

$$f(x^*) \le f(x(\bar\beta)) \le \bar\beta f(x^\circ) + (1-\bar\beta) f(x^*)$$

da cui segue ma $f(x^\circ) < f(x^*)$! $\square$

Dimostrazione – continua

Consideriamo il segmento $[x^*, x^\circ]$ che congiunge x^* con x°. Per convessità, risulta $[x^*, x^\circ] \subset \mathcal{F}$ e

$$f(x(\beta)) \leq \beta f(x^\circ) + (1-\beta) f(x^*), \quad \forall\ \beta \in [0,1]$$

Deve esistere un valore $\bar{\beta} \in (0,1)$ tale che il punto

$$x(\bar{\beta}) = \bar{\beta} x^\circ + (1-\bar{\beta}) x^* \in [x^*, x^\circ]$$

appartiene a $\mathcal{F} \cap \mathcal{B}(x^*; \epsilon)$ e quindi

$$f(x^*) \leq f(x(\bar{\beta})) \leq \bar{\beta} f(x^\circ) + (1-\bar{\beta}) f(x^*)$$

da cui segue ma $f(x^\circ) < f(x^*)$!

Dimostrazione – continua

Consideriamo il segmento $[x^*, x^\circ]$ che congiunge x^* con x°. Per convessità, risulta $[x^*, x^\circ] \subset \mathcal{F}$ e

$$f(x(\beta)) \le \beta f(x^\circ) + (1-\beta) f(x^*), \quad \forall\ \beta \in [0,1]$$

Deve esistere un valore $\bar\beta \in (0,1)$ tale che il punto

$$x(\bar\beta) = \bar\beta x^\circ + (1-\bar\beta)x^* \in [x^*, x^\circ]$$

appartiene a $\mathcal{F} \cap \mathcal{B}(x^*;\epsilon)$ e quindi

$$f(x^*) \le f(x(\bar\beta)) \le \bar\beta f(x^\circ) + (1-\bar\beta) f(x^*)$$

da cui segue ma $f(x^\circ) < f(x^*)$! □

Dimostrazione – continua

Consideriamo il segmento $[x^*, x^\circ]$ che congiunge x^* con x°. Per convessità, risulta $[x^*, x^\circ] \subset \mathcal{F}$ e

$$f(x(\beta)) \leq \beta f(x^\circ) + (1-\beta) f(x^*), \quad \forall\, \beta \in [0,1]$$

Deve esistere un valore $\bar{\beta} \in (0,1)$ tale che il punto

$$x(\bar{\beta}) = \bar{\beta} x^\circ + (1-\bar{\beta}) x^* \in [x^*, x^\circ]$$

appartiene a $\mathcal{F} \cap \mathcal{B}(x^*; \epsilon)$ e quindi

$$f(x^*) \leq f(x(\bar{\beta})) \leq \bar{\beta} f(x^\circ) + (1-\bar{\beta}) f(x^*)$$

da cui segue

$$f(x^*) \leq \bar{\beta} f(x^\circ) + (1-\bar{\beta}) f(x^*)$$

ma $f(x^\circ) < f(x^*)$! $\square$

Dimostrazione – continua

Consideriamo il segmento $[x^*, x^\circ]$ che congiunge x^* con x°. Per convessità, risulta $[x^*, x^\circ] \subset \mathcal{F}$ e

$$f(x(\beta)) \leq \beta f(x^\circ) + (1-\beta) f(x^*), \quad \forall\, \beta \in [0,1]$$

Deve esistere un valore $\bar{\beta} \in (0,1)$ tale che il punto

$$x(\bar{\beta}) = \bar{\beta} x^\circ + (1-\bar{\beta}) x^* \in [x^*, x^\circ]$$

appartiene a $\mathcal{F} \cap \mathcal{B}(x^*; \epsilon)$ e quindi

$$f(x^*) \leq f(x(\bar{\beta})) \leq \bar{\beta} f(x^\circ) + (1-\bar{\beta}) f(x^*)$$

da cui segue

$$0 \leq \bar{\beta} f(x^\circ) - \bar{\beta} f(x^*)$$

ma $f(x^\circ) < f(x^*)$! $\square$

Dimostrazione – continua

Consideriamo il segmento $[x^*, x^\circ]$ che congiunge x^* con x°. Per convessità, risulta $[x^*, x^\circ] \subset \mathcal{F}$ e

$$f(x(\beta)) \leq \beta f(x^\circ) + (1-\beta) f(x^*), \quad \forall\ \beta \in [0,1]$$

Deve esistere un valore $\bar{\beta} \in (0,1)$ tale che il punto

$$x(\bar{\beta}) = \bar{\beta} x^\circ + (1-\bar{\beta}) x^* \in [x^*, x^\circ]$$

appartiene a $\mathcal{F} \cap \mathcal{B}(x^*; \epsilon)$ e quindi

$$f(x^*) \leq f(x(\bar{\beta})) \leq \bar{\beta} f(x^\circ) + (1-\bar{\beta}) f(x^*)$$

da cui segue

$$\bar{\beta} f(x^*) \leq \bar{\beta} f(x^\circ)$$

ma $f(x^\circ) < f(x^*)$! $\square$

Dimostrazione – continua

Consideriamo il segmento $[x^*, x^\circ]$ che congiunge x^* con x°. Per convessità, risulta $[x^*, x^\circ] \subset \mathcal{F}$ e

$$f(x(\beta)) \leq \beta f(x^\circ) + (1-\beta) f(x^*), \quad \forall\, \beta \in [0,1]$$

Deve esistere un valore $\bar{\beta} \in (0,1)$ tale che il punto

$$x(\bar{\beta}) = \bar{\beta} x^\circ + (1-\bar{\beta}) x^* \in [x^*, x^\circ]$$

appartiene a $\mathcal{F} \cap \mathcal{B}(x^*; \epsilon)$ e quindi

$$f(x^*) \leq f(x(\bar{\beta})) \leq \bar{\beta} f(x^\circ) + (1-\bar{\beta}) f(x^*)$$

da cui segue

$$f(x^*) \leq f(x^\circ)$$

ma $f(x^\circ) < f(x^*)$! $\square$

Dimostrazione – continua

Consideriamo il segmento $[x^*, x^\circ]$ che congiunge x^* con x°. Per convessità, risulta $[x^*, x^\circ] \subset \mathcal{F}$ e

$$f(x(\beta)) \leq \beta f(x^\circ) + (1-\beta) f(x^*), \quad \forall\, \beta \in [0,1]$$

Deve esistere un valore $\bar{\beta} \in (0,1)$ tale che il punto

$$x(\bar{\beta}) = \bar{\beta} x^\circ + (1-\bar{\beta}) x^* \in [x^*, x^\circ]$$

appartiene a $\mathcal{F} \cap \mathcal{B}(x^*; \epsilon)$ e quindi

$$f(x^*) \leq f(x(\bar{\beta})) \leq \bar{\beta} f(x^\circ) + (1-\bar{\beta}) f(x^*)$$

da cui segue

$$f(x^*) \leq f(x^\circ)$$

ma $f(x^\circ) < f(x^*)$! $\square$

Condizioni sufficienti di ottimalità

Supponiamo:

- f, g continuamente differenziabili e **convesse**
- h **affini**, e.g. $h_j(x) = a^\top x - b = 0$

Teorema (Condizione Sufficiente)

Se $x^* \in \mathcal{F}$ è un **punto di KKT** *allora x^* è un* **minimo globale** *del problema*

Dim. x^* KKT quindi, ci sono dei moltiplicatori $\lambda^* \geq \mathbf{0}$, μ^* tali che

$$\nabla f(x^*) + \nabla g(x^*)\lambda^* + \nabla h(x^*)\mu^* = \mathbf{0}, \qquad g(x^*)^\top \lambda^* = 0$$

Condizioni sufficienti di ottimalità

Supponiamo:

- f, g continuamente differenziabili e **convesse**
- h **affini**, e.g. $h_j(x) = a^\top x - b = 0$

Teorema (Condizione Sufficiente)

Se $x^* \in \mathcal{F}$ è un **punto di KKT** *allora x^* è un* **minimo globale** *del problema*

Dim. x^* KKT quindi, ci sono dei moltiplicatori $\lambda^* \geq \mathbf{0}$, μ^* tali che

$$\nabla f(x^*) + \nabla g(x^*)\lambda^* + \nabla h(x^*)\mu^* = \mathbf{0}, \qquad g(x^*)^\top \lambda^* = 0$$

Condizioni sufficienti di ottimalità

Supponiamo:

- f, g continuamente differenziabili e **convesse**
- h **affini**, e.g. $h_j(x) = a^\top x - b = 0$

Teorema (Condizione Sufficiente)

Se $x^* \in \mathcal{F}$ è un **punto di KKT** *allora x^* è un* **minimo globale** *del problema*

Dim. x^* KKT quindi, ci sono dei moltiplicatori $\lambda^* \geq \mathbf{0}$, μ^* tali che

$$\nabla f(x^*) + \nabla g(x^*)\lambda^* + \nabla h(x^*)\mu^* = \mathbf{0}, \qquad g(x^*)^\top \lambda^* = 0$$

Dimostrazione – continua

Per ogni punto ammissibile $x \in \mathcal{F}$, $(g(x) \leq 0, h(x) = 0)$ abbiamo

$$f(x) \geq f(x) + g(x)^\top \lambda^* + h(x)^\top \mu^*$$

Dalla linearità di $h(x)$ segue che

$$h(x) = h(x^*) + \nabla h(x^*)^\top (x - x^*)$$

Dalla convessità di f e delle g_i segue invece che

$$\begin{aligned} f(x) &\geq f(x^*) + \nabla f(x^*)^\top (x - x^*) \\ g_i(x) &\geq g_i(x^*) + \nabla g_i(x^*)^\top (x - x^*), \quad i = 1, \dots, m \end{aligned}$$

e siccome $\lambda^* \geq \mathbf{0}$,

Dimostrazione – continua

Per ogni punto ammissibile $x \in \mathcal{F}$, $(g(x) \leq 0, h(x) = 0)$ abbiamo

$$f(x) \geq f(x) + g(x)^\top \lambda^* + h(x)^\top \mu^*$$

Dalla linearità di $h(x)$ segue che

$$h(x) = h(x^*) + \nabla h(x^*)^\top (x - x^*)$$

Dalla convessità di f e delle g_i segue invece che

$$\begin{aligned} f(x) &\geq f(x^*) + \nabla f(x^*)^\top (x - x^*) \\ g_i(x) &\geq g_i(x^*) + \nabla g_i(x^*)^\top (x - x^*), \quad i = 1, \dots, m \end{aligned}$$

e siccome $\lambda^* \geq \mathbf{0}$,

Dimostrazione – continua

Per ogni punto ammissibile $x \in \mathcal{F}$, $(g(x) \leq 0, h(x) = 0)$ abbiamo

$$f(x) \geq f(x) + g(x)^\top \lambda^* + h(x)^\top \mu^*$$

Dalla linearità di $h(x)$ segue che

$$h(x) = h(x^*) + \nabla h(x^*)^\top (x - x^*)$$

Dalla convessità di f e delle g_i segue invece che

$$\begin{array}{rcl} f(x) & \geq & f(x^*) + \nabla f(x^*)^\top (x - x^*) \\ g_i(x) & \geq & g_i(x^*) + \nabla g_i(x^*)^\top (x - x^*), \quad i = 1, \dots, m \end{array}$$

e siccome $\lambda^* \geq \mathbf{0}$,

Dimostrazione – continua

Per ogni punto ammissibile $x \in \mathcal{F}$, $(g(x) \leq 0, h(x) = 0)$ abbiamo

$$f(x) \geq f(x) + g(x)^\top \lambda^* + h(x)^\top \mu^*$$

Dalla linearità di $h(x)$ segue che

$$h(x) = h(x^*) + \nabla h(x^*)^\top (x - x^*)$$

Dalla convessità di f e delle g_i segue invece che

$$\begin{array}{rcl} f(x) & \geq & f(x^*) + \nabla f(x^*)^\top (x - x^*) \\ g_i(x) & \geq & g_i(x^*) + \nabla g_i(x^*)^\top (x - x^*), \quad i = 1, \dots, m \end{array}$$

e siccome $\lambda^* \geq \mathbf{0}$,

$$\lambda_i^* g_i(x) \geq \lambda_i^* g_i(x^*) + \lambda_i^* \nabla g_i(x^*)^\top (x - x^*), \quad i = 1, \dots, m$$

Dimostrazione – continua

Per ogni punto ammissibile $x \in \mathcal{F}$, $(g(x) \leq 0, h(x) = 0)$ abbiamo

$$f(x) \geq f(x) + g(x)^\top \lambda^* + h(x)^\top \mu^*$$

Dalla linearità di $h(x)$ segue che

$$h(x) = h(x^*) + \nabla h(x^*)^\top (x - x^*)$$

Dalla convessità di f e delle g_i segue invece che

$$\begin{array}{rcl} f(x) & \geq & f(x^*) + \nabla f(x^*)^\top (x - x^*) \\ g_i(x) & \geq & g_i(x^*) + \nabla g_i(x^*)^\top (x - x^*), \quad i = 1, \ldots, m \end{array}$$

e siccome $\lambda^* \geq \mathbf{0}$,

$$(\lambda^*)^\top g(x) \geq (\lambda^*)^\top g(x^*) + (\lambda^*)^\top \nabla g(x^*)^\top (x - x^*)$$

Dimostrazione – continua

Quindi, possiamo scrivere

$$
\begin{aligned}
f(x) &\geq f(x) + g(x)^\top \lambda^* + h(x)^\top \mu^* \\
&\geq f(x^*) + \nabla f(x^*)^\top (x - x^*) + \underbrace{(\lambda^*)^\top g(x^*)}_{0} + (\lambda^*)^\top \nabla g(x^*)^\top (x - x^*) + \\
&\quad \underbrace{(\mu^*)^\top h(x^*)}_{0} + (\mu^*)^\top \nabla h(x^*)^\top (x - x^*) \\
&= f(x^*) + \nabla f(x^*)^\top (x - x^*) + (\lambda^*)^\top \nabla g(x^*)^\top (x - x^*) + \\
&\quad (\mu^*)^\top \nabla h(x^*)^\top (x - x^*) \\
&= f(x^*) + \underbrace{(\nabla f(x^*) + \nabla g(x^*)\lambda^* + \nabla h(x^*)\mu^*)^\top}_{\nabla_x L^* = 0} (x - x^*) \\
&= f(x^*)
\end{aligned}
$$

Dimostrazione – continua

Quindi, possiamo scrivere

$$
\begin{aligned}
f(x) &\geq f(x) + g(x)^\top \lambda^* + h(x)^\top \mu^* \\
&\geq f(x^*) + \nabla f(x^*)^\top (x - x^*) + \underbrace{(\lambda^*)^\top g(x^*)}_{0} + (\lambda^*)^\top \nabla g(x^*)^\top (x - x^*) + \\
&\quad \underbrace{(\mu^*)^\top h(x^*)}_{0} + (\mu^*)^\top \nabla h(x^*)^\top (x - x^*) \\
&= f(x^*) + \nabla f(x^*)^\top (x - x^*) + (\lambda^*)^\top \nabla g(x^*)^\top (x - x^*) + \\
&\quad (\mu^*)^\top \nabla h(x^*)^\top (x - x^*) \\
&= f(x^*) + \underbrace{(\nabla f(x^*) + \nabla g(x^*)\lambda^* + \nabla h(x^*)\mu^*)^\top}_{\nabla_x L^* = \mathbf{0}} (x - x^*) \\
&= f(x^*)
\end{aligned}
$$

Dimostrazione – continua

Quindi, possiamo scrivere

$$\begin{aligned}
f(x) &\geq f(x) + g(x)^\top \lambda^* + h(x)^\top \mu^* \\
&\geq f(x^*) + \nabla f(x^*)^\top (x - x^*) + \underbrace{(\lambda^*)^\top g(x^*)}_{0} + (\lambda^*)^\top \nabla g(x^*)^\top (x - x^*) + \\
&\quad \underbrace{(\mu^*)^\top h(x^*)}_{0} + (\mu^*)^\top \nabla h(x^*)^\top (x - x^*) \\
&= f(x^*) + \nabla f(x^*)^\top (x - x^*) + (\lambda^*)^\top \nabla g(x^*)^\top (x - x^*) + \\
&\quad (\mu^*)^\top \nabla h(x^*)^\top (x - x^*) \\
&= f(x^*) + \underbrace{(\nabla f(x^*) + \nabla g(x^*)\lambda^* + \nabla h(x^*)\mu^*)^\top}_{\nabla_x L^* = \mathbf{0}} (x - x^*) \\
&= f(x^*)
\end{aligned}$$

Dimostrazione – continua

Quindi, possiamo scrivere

$$\begin{aligned}
f(x) &\geq f(x) + g(x)^\top \lambda^* + h(x)^\top \mu^* \\
&\geq f(x^*) + \nabla f(x^*)^\top (x - x^*) + \underbrace{(\lambda^*)^\top g(x^*)}_{0} + (\lambda^*)^\top \nabla g(x^*)^\top (x - x^*) + \\
&\quad \underbrace{(\mu^*)^\top h(x^*)}_{0} + (\mu^*)^\top \nabla h(x^*)^\top (x - x^*) \\
&= f(x^*) + \nabla f(x^*)^\top (x - x^*) + (\lambda^*)^\top \nabla g(x^*)^\top (x - x^*) + \\
&\quad (\mu^*)^\top \nabla h(x^*)^\top (x - x^*) \\
&= f(x^*) + \underbrace{(\nabla f(x^*) + \nabla g(x^*)\lambda^* + \nabla h(x^*)\mu^*)^\top}_{\nabla_x L^* = \mathbf{0}} (x - x^*) \\
&= f(x^*)
\end{aligned}$$

Dimostrazione – continua

Quindi, possiamo scrivere

$$
\begin{aligned}
f(x) &\geq f(x) + g(x)^\top \lambda^* + h(x)^\top \mu^* \\
&\geq f(x^*) + \nabla f(x^*)^\top (x - x^*) + \underbrace{(\lambda^*)^\top g(x^*)}_{0} + (\lambda^*)^\top \nabla g(x^*)^\top (x - x^*) + \\
&\quad \underbrace{(\mu^*)^\top h(x^*)}_{0} + (\mu^*)^\top \nabla h(x^*)^\top (x - x^*) \\
&= f(x^*) + \nabla f(x^*)^\top (x - x^*) + (\lambda^*)^\top \nabla g(x^*)^\top (x - x^*) + \\
&\quad (\mu^*)^\top \nabla h(x^*)^\top (x - x^*) \\
&= f(x^*) + \underbrace{(\nabla f(x^*) + \nabla g(x^*)\lambda^* + \nabla h(x^*)\mu^*)^\top}_{\nabla_x L^* = \mathbf{0}} (x - x^*) \\
&= f(x^*)
\end{aligned}
$$

Condizioni necessarie di ottimalità

Proposizione (Ottimizzazione Non vincolata)

Sia x^* *un minimo locale* **non vincolato** *di* f*. Allora*

$$d^\top \nabla^2 f(x^*) d \geq 0, \quad \forall\, d \in \mathbb{R}^n$$

Condizioni necessarie di ottimalità

Proposizione

Supponiamo:

- f, g, h **due volte** *continuamente differenziabili*
- x^* **minimo locale regolare** *del problema vincolato*

Allora, x^* è un punto di KKT con moltiplicatori λ^*, μ^*. Inoltre risulta

$$d^\top \nabla_x^2 L(x^*, \lambda^*, \mu^*) d \geq 0, \quad \forall\, d: \; M^* d = \mathbf{0}$$

dove

$$M^* = \left(\begin{array}{c} \nabla h_j(x^*)^\top \\ \nabla g_i(x^*)^\top \end{array} \right)_{\Big|\, j=1,\dots,p,\, i \in I_0(x^*)}$$

Condizioni necessarie di ottimalità

Proposizione

Supponiamo:

- f, g, h **due volte** *continuamente differenziabili*
- x^* **minimo locale regolare** *del problema vincolato*

Allora, x^* è un punto di KKT con moltiplicatori λ^*, μ^*. Inoltre risulta

$$d^\top \nabla_x^2 L(x^*, \lambda^*, \mu^*) d \geq 0, \quad \forall\ d:\ M^* d = \mathbf{0}$$

dove

$$M^* = \left(\begin{array}{c} \nabla h_j(x^*)^\top \\ \nabla g_i(x^*)^\top \end{array} \right)_{\Big|\, j=1,\dots,p,\, i\in I_0(x^*)}$$

Condizioni necessarie di ottimalità

Proposizione

Supponiamo:

- f, g, h **due volte** *continuamente differenziabili*
- x^* **minimo locale regolare** *del problema vincolato*

Allora, x^* è un punto di KKT con moltiplicatori λ^*, μ^*. Inoltre risulta

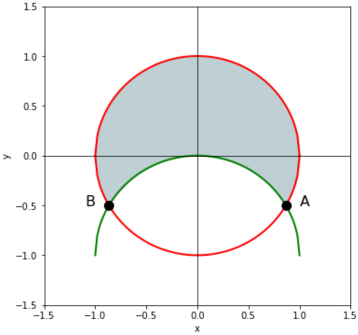
$$d^\top \nabla_x^2 L(x^*, \lambda^*, \mu^*) d \geq 0, \quad \forall\ d:\ M^* d = \mathbf{0}$$

dove

$$M^* = \left(\begin{array}{c} \nabla h_j(x^*)^\top \\ \nabla g_i(x^*)^\top \end{array} \right)_{\Big|\, j=1,\dots,p,\, i \in I_0(x^*)}$$

Esempio

$$\begin{aligned} \min\ & x + y \\ s.t.\ & x^2 + y^2 \leq 1 \\ & x^2 + (y+1)^2 \geq 1 \end{aligned}$$



$$A = (\sqrt{3}/2, -1/2)^\top$$
$$B = (-\sqrt{3}/2, -1/2)^\top$$

Esempio

Il **punto A** è un punto di KKT con moltiplicatori

$$\lambda_1^* = \frac{3-\sqrt{3}}{6}, \quad \lambda_2^* = \frac{3+\sqrt{3}}{6}$$

risulta

$$\nabla_x^2 L(A, \lambda^*) = \begin{pmatrix} \frac{3-\sqrt{3}}{3} - \frac{3+\sqrt{3}}{3} & 0 \\ 0 & \frac{3-\sqrt{3}}{3} - \frac{3+\sqrt{3}}{3} \end{pmatrix} = \begin{pmatrix} -\frac{2\sqrt{3}}{3} & 0 \\ 0 & -\frac{2\sqrt{3}}{3} \end{pmatrix}$$

$$M^* = \begin{pmatrix} \sqrt{3} & -1 \\ -\sqrt{3} & -1 \end{pmatrix}$$

Percui, $M^* d = 0$ implica $d = 0$ e la C.N. del II ordine è soddisfatta in A

Esempio

Il **punto B** è un punto di KKT con moltiplicatori

$$\lambda_1^* = \frac{3+\sqrt{3}}{6}, \quad \lambda_2^* = \frac{3-\sqrt{3}}{6}$$

risulta

$$\nabla_x^2 L(B,\lambda^*) = \begin{pmatrix} \frac{3+\sqrt{3}}{3} - \frac{3-\sqrt{3}}{3} & 0 \\ 0 & \frac{3+\sqrt{3}}{3} - \frac{3-\sqrt{3}}{3} \end{pmatrix} = \begin{pmatrix} \frac{2\sqrt{3}}{3} & 0 \\ 0 & \frac{2\sqrt{3}}{3} \end{pmatrix}$$

che è definita positiva. Percui, la C.N. del II ordine è soddisfatta anche in B

Condizioni sufficienti di ottimalità

Proposizione (Ottimizzazione Non vincolata)

Sia $x^* \in \mathbb{R}^n$ tale che

- $\nabla f(x^*) = \mathbf{0}$
- $\nabla^2 f(x^*)$ *è definita positiva, i.e.*

$$d^\top \nabla^2 f(x^*) d > 0, \quad \forall\, d \in \mathbb{R}^n \ (d \neq \mathbf{0})$$

Allora x^* è minimo locale stretto di f

$x^* = 0$ è minimo locale stretto di $f(x) = x^4$ ma, essendo

$$\nabla^2 f(0) = \mathbf{0},$$

x^* non verifica le cond. sufficienti

Condizioni sufficienti di ottimalità

Teorema (Condizione Sufficiente Forte)

Supponiamo:

- f, g, h **due volte** *continuamente differenziabili*
- x^* **punto di KKT** *con molt.* λ^*, μ^*

Se $d^\top \nabla^2_{xx} L(x^*, \lambda^*, \mu^*) d > 0$, *per ogni* $d \in Y^*$, $d \neq 0$, *ove*

$$
\begin{aligned}
Y^* \quad = \quad \Big\{ d \in \mathbb{R}^n : \nabla h_j(x^*)^\top d = 0,\ j = 1, \dots, p \\
\nabla g_i(x^*)^\top d \le 0,\ i \in I_0(x^*), \\
\nabla g_i(x^*)^\top d = 0,\ i \in I_0(x^*), \lambda_i^* > 0 \Big\}
\end{aligned}
$$

allora x^* *è un minimo locale stretto del problema*

Condizioni sufficienti di ottimalità

Teorema (Condizione Sufficiente Forte)

Supponiamo:

- f, g, h **due volte** *continuamente differenziabili*
- x^* **punto di KKT** *con molt.* λ^*, μ^*

Se $d^\top \nabla^2_{xx} L(x^*, \lambda^*, \mu^*) d > 0$, *per ogni* $d \in Y^*$, $d \neq 0$, *ove*

$$
\begin{aligned}
Y^* \quad = \quad \Big\{ d \in \mathbb{R}^n : & \nabla h_j(x^*)^\top d = 0,\ j = 1, \ldots, p \\
& \nabla g_i(x^*)^\top d \le 0,\ i \in I_0(x^*), \\
& \nabla g_i(x^*)^\top d = 0,\ i \in I_0(x^*), \lambda_i^* > 0 \Big\}
\end{aligned}
$$

allora x^* *è un minimo locale stretto del problema*

Condizioni sufficienti di ottimalità

Teorema (Condizione Sufficiente Forte)

Supponiamo:

- f, g, h **due volte** *continuamente differenziabili*
- x^* **punto di KKT** *con molt.* λ^*, μ^*

Se $d^\top \nabla^2_{xx} L(x^*, \lambda^*, \mu^*) d > 0$, *per ogni* $d \in Y^*$, $d \neq 0$, *ove*

$$
\begin{aligned}
Y^* \;=\; \Big\{ d \in \mathbb{R}^n : \; & \nabla h_j(x^*)^\top d = 0, \; j = 1, \dots, p \\
& \nabla g_i(x^*)^\top d \leq 0, \; i \in I_0(x^*), \\
& \nabla g_i(x^*)^\top d = 0, \; i \in I_0(x^*), \lambda_i^* > 0 \Big\}
\end{aligned}
$$

allora x^* *è un minimo locale stretto del problema*

Condizioni sufficienti di ottimalità

Teorema (Condizione Sufficiente Forte)

Supponiamo:

- f, g, h **due volte** *continuamente differenziabili*
- x^* **punto di KKT** *con molt.* λ^*, μ^*

Se $d^\top \nabla^2_{xx} L(x^*, \lambda^*, \mu^*) d > 0$, *per ogni* $d:\ M^* d = \mathbf{0},\ d \neq 0$, *ove*

$$M^* = \left. \begin{pmatrix} \nabla h_j(x^*)^\top \\ \nabla g_i(x^*)^\top \end{pmatrix} \right|_{j=1,\dots,p,\ i \in I_0^+(x^*)}$$

allora x^* *è un minimo locale stretto del problema*