

Ottimizzazione Non Lineare

G. Liuzzi¹

Lunedì 17 Dicembre 2018

¹Istituto di Analisi dei Sistemi ed Informatica IASI - CNR



Condizioni di ottimalità – vinc. di uguaglianza

$$\begin{array}{ll} \min_x & f(x) \\ \text{c.v.} & h(x) = 0 \end{array} \quad (P)$$

Supponiamo:

- $h : \mathbb{R}^n \rightarrow \mathbb{R}$ (**un solo** vincolo di uguaglianza)
- f, h continuamente differenziabili
- x^* minimo locale di (P) tale che $\nabla h(x^*) \neq \mathbf{0}$

Facciamo vedere che allora $\nabla f(x^*)$ e $\nabla h(x^*)$ devono essere collineari cioè, deve esistere $\sigma \in \mathbb{R}$ tale che

$$\nabla f(x^*) = \sigma \nabla h(x^*)$$

È banale se $\nabla f(x^*) = \mathbf{0}$, quindi supponiamo $\nabla f^* \neq \mathbf{0}$



Condizioni di ottimalità – vinc. di uguaglianza (segue)

Dalla definizione di minimo locale segue che esiste $\epsilon > 0$ tale che

$$f(x) \geq f(x^*),$$

per ogni $x \in \mathcal{B}(x^*; \epsilon) = \mathcal{B}$ e tale che $h(x) = 0$

La curva di livello $f(x) = f(x^*) = c$ ci permette di partizionare \mathcal{B} in

$$\mathcal{B}^+ \equiv \{x \in \mathcal{B} : f(x) > c\}$$

$$\mathcal{B}^- \equiv \{x \in \mathcal{B} : f(x) < c\}$$

Se ∇f^* e ∇h^* (non nulli) non fossero collineari, allora la curva $h(x) = 0$ intersecherebbe la curva $f(x) = c$ in x^* .

Quindi esisterebbero punti x tali che $x \in \mathcal{B}^-$ e $h(x) = 0$. Per tali punti risulterebbe $f(x) < c = f(x^*)$ contraddicendo l'ipotesi che x^* è minimo locale di (P).



Condizioni di ottimalità – vinc. di uguaglianza (segue)

x^* minimo locale t.c. $\nabla h(x^*) \neq 0$, allora:

Teorema (Kaursh, 1939, Kuhn-Tucker, 1951)

Esiste un moltiplicatore μ^ tale che:*

$$\nabla f(x^*) + \mu^* \nabla h(x^*) = 0,$$

x^* minimo locale, allora:

Teorema (Fritz-John, 1948)

Esistono moltiplicatori $\lambda_0^ \geq 0$ e μ^* (non entrambi nulli) tale che:*

$$\lambda_0^* \nabla f(x^*) + \mu^* \nabla h(x^*) = 0,$$



Condizioni di ottimalità – vinc. di uguaglianza

$$\begin{array}{ll} \min_x & f(x) \\ \text{c.v.} & h(x) = 0 \end{array} \quad (P)$$

supponiamo:

- $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$ ($1 \leq p < n$)
- f, h continuamente differenziabili
- x^* minimo locale di (P) tale che $\{\nabla h_i(x^*), i = 1, \dots, p\}$ lin.indipendenti

Facciamo vedere che allora $\nabla f(x^*)$ è combinazione lineare di $\nabla h_i(x^*), i = 1, \dots, p$.

Il caso $\nabla f(x^*) = 0$ è banale, quindi supponiamo $\nabla f(x^*) \neq 0$



Teorema (T1)

Teorema (M.R.Hestenes, 1975)

Siano \bar{x} e d tali che

- $h(\bar{x}) = 0$;
- $\nabla h_i(\bar{x})$, $i = 1, \dots, p$ lin. indipendenti;
- $\nabla h_i(\bar{x})^\top d = 0$.

Allora, è possibile definire una curva $x(t) \in C^1$, $-\delta \leq t \leq \delta$, tale che

- $h(x(t)) = 0$;
- $x(0) = \bar{x}$;
- $\dot{x}(0) = d$.



Condizioni di ottimalità – vinc. di uguaglianza (segue)

Per il fatto che $p < n$, segue che il sistema omogeneo

$$\nabla h_i(x^*)^\top d = 0, \quad i = 1, \dots, p$$

(è sottodimensionato e) ammette (infinite) soluzioni $d \neq 0$.

Tali soluzioni sono i vettori “tangenti” alla superficie $h(x) = 0$ in x^* .

Il Teorema (T1), comunque scelta una soluzione d (tangente), esiste una curva $x(t) \in C^1$, $-\delta < t < \delta$ sulla superficie tale che

$$h(x(t)) = 0. \quad x(0) = x^*, \quad \dot{x}(0) = d$$

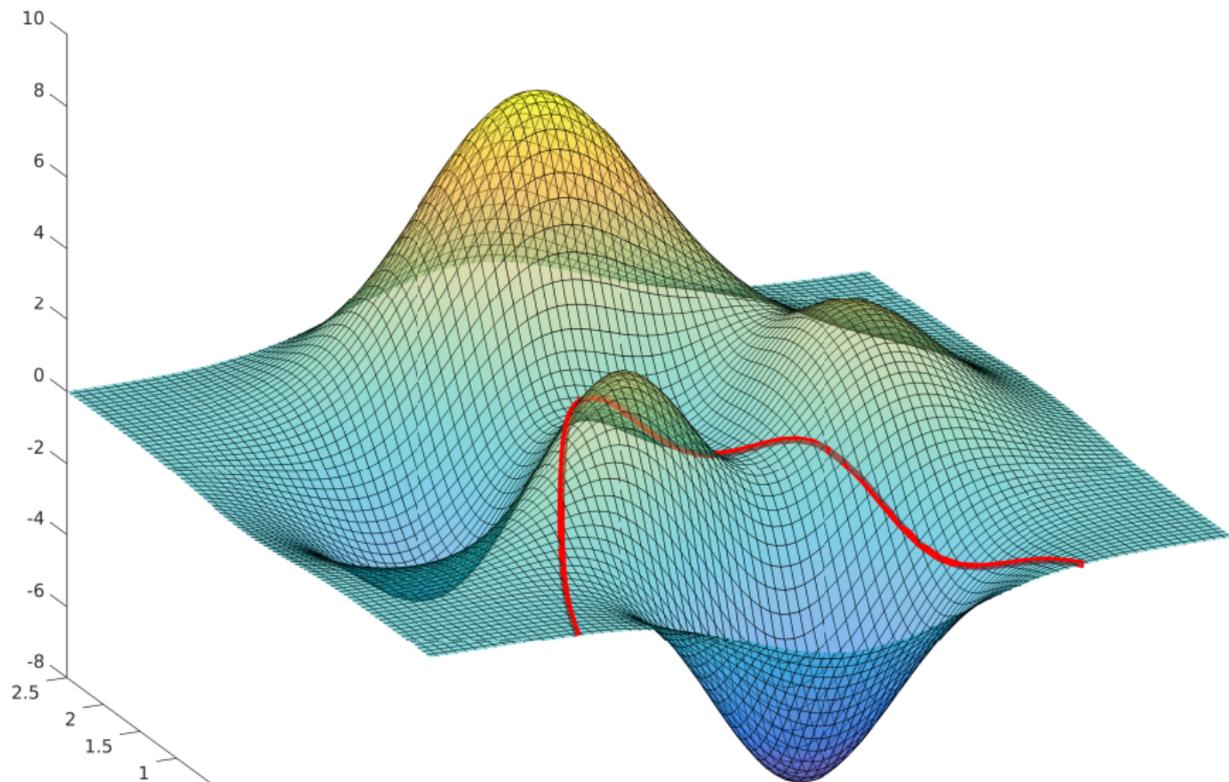
Lungo la curva $x(t)$ la funzione $f(x)$ diviene

$$\psi(t) = f(x(t))$$

con $\psi(0) = f(x^*)$ e $\psi'(0) = \nabla f(x^*)^\top \dot{x}(0) = \nabla f(x^*)^\top d$.



Figura



Condizioni di ottimalità – vinc. di uguaglianza (segue)

Dato che x^* è minimo locale, la funzione $\psi(t)$ deve anch'essa avere un minimo locale in $t = 0$ e quindi $\psi'(0) = 0$

$$\psi'(0) = \nabla f(x^*)^\top d = 0$$

Questo vuol dire che tutti i vettori d tangenti alla superficie $h(x) = 0$ in x^* sono anche ortogonali al gradiente di f in x^*

Quindi i due sistemi di equazioni lineari omogenei

$$\nabla h_i(x^*)^\top d = 0, \quad i = 1, \dots, p \quad \text{e} \quad \begin{array}{l} \nabla h_i(x^*)^\top d = 0, \quad i = 1, \dots, p \\ \nabla f(x^*)^\top d = 0 \end{array}$$

devono avere lo stesso insieme di soluzioni. In altri termini, l'equazione $\nabla f(x^*)^\top d$ deve essere ridondante, cioè devono esistere numeri σ_i , $i = 1, \dots, p$ tali che

$$\nabla f(x^*) = \sum_{i=1}^p \sigma_i \nabla h_i(x^*).$$



Condizioni di ottimalità – vinc. di uguaglianza (segue)

x^* minimo locale t.c. $\nabla h_i(x^*)$ lin. indipendenti, allora:

Teorema (Kaursh, 1939, Kuhn-Tucker, 1951)

Esistono moltiplicatori μ_i^ tali che:*

$$\nabla f(x^*) + \sum_{i=1}^p \mu_i^* \nabla h_i(x^*) = 0,$$

x^* minimo locale, allora:

Teorema (Fritz-John, 1948)

Esistono moltiplicatori $\lambda_0^ \geq 0$ e μ_i^* (non tutti nulli) tali che:*

$$\lambda_0^* \nabla f(x^*) + \sum_{i=1}^p \mu_i^* \nabla h_i(x^*) = 0,$$



Condizioni di ottimalità

Supponiamo:

- f, g, h continuamente differenziabili
- di conoscere un punto di **minimo locale** x^* del problema (P_0)
- che in x^* i vincoli siano **regolari**

e.g. $\left\{ \nabla h_j(x^*), j = 1, \dots, p, \nabla g_i(x^*), i \in I_0(x^*) \right\}$ lin. indep.

Allora

Teorema (Kaursh, 1939, Kuhn-Tucker, 1951)

Esistono dei moltiplicatori $\lambda_1^, \dots, \lambda_m^* \geq 0, \mu_1^*, \dots, \mu_p^*$ tali che:*

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^r \mu_j^* \nabla h_j(x^*) = 0,$$

$$\lambda_i^* g_i(x^*) = 0, \quad \text{per ogni } i = 1, \dots, m.$$



Esempio

$$\begin{array}{ll} \min_x & x_1 + x_2 \\ \text{c.v} & x_1^2 + x_2^2 - 2 = 0 \end{array}$$

$$\begin{array}{ll} \min_x & x_1 + x_2 \\ \text{c.v} & x_1^2 + x_2^2 - 2 \leq 0 \end{array}$$

$$\begin{array}{ll} \min_x & x_1 + x_2 \\ \text{c.v} & x_1^2 + x_2^2 - 2 \leq 0, \quad x_2 \geq 0 \end{array}$$

$$\begin{array}{ll} \min_x & x_1 + x_2 \\ \text{c.v} & x_2 - x_1^3 \leq 0, \quad x_2 \geq 0 \end{array}$$

