

DISAMINA DI METODI DI OTTIMIZZAZIONE FONDATA SULL'IMPIEGO DEI MOLTIPLICATO- RI DI LAGRANGE

Corso Metodi Matematici

23 settembre 2018

Pietro Pennestrì

OTTIMIZZAZIONE CON VINCOLI DI UGUAGLIANZA

Siano:

$$f, g : \mathbb{R}^n \rightarrow \mathbb{R}$$

$$f, g \in C^1(A)$$

dove A è un aperto di \mathbb{R}^n

Minimizzare $f = f(\mathbf{x})$, con $\mathbf{x} \in \mathbb{R}^n$

soggetta a

$$g_j(\mathbf{x}) = 0, \quad j = 1, 2, \dots, m.$$

Nel seguito assumeremo sempre $m < n$, in caso contrario il problema diventa sovradefinito e nella maggior parte dei casi non avrà soluzione.

L'idea alla base del *metodo delle variazioni vincolate* è quella di determinare un'espressione in forma chiusa del differenziale df per tutti i punti per cui $g_i(\mathbf{x}) = 0$, ($i = 1, 2, \dots, m$) è soddisfatta. Per fissare le idee il metodo verrà sviluppato per il caso bidimensionale $n = 2$ e $m = 1$.

Minimizzare $f(x_1, x_2)$ soggetta a $g(x_1, x_2) = 0$

Minimizzare $f(x_1, x_2)$ soggetta a $g(x_1, x_2) = 0$

In generale, affinché ad x^* corrisponda un minimo di $f(\mathbf{x})$ è necessario sia:

$$\frac{\partial f(\mathbf{x}^*)}{\partial x_1} = \frac{\partial f(\mathbf{x}^*)}{\partial x_2} = \dots = \frac{\partial f(\mathbf{x}^*)}{\partial x_n} = 0$$

Minimizzare $f(x_1, x_2)$ soggetta a $g(x_1, x_2) = 0$

In generale, affinché ad x^* corrisponda un minimo di $f(\mathbf{x})$ è necessario sia:

$$\frac{\partial f(\mathbf{x}^*)}{\partial x_1} = \frac{\partial f(\mathbf{x}^*)}{\partial x_2} = \dots = \frac{\partial f(\mathbf{x}^*)}{\partial x_n} = 0$$

ovvero:

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \dots + \frac{\partial f}{\partial x_n} dx_n = 0 \quad (1)$$

Poiché il vincolo impone $g(x_1^*, x_2^*) = 0$, possiamo ragionevolmente assumere

$$g(x_1^* + dx_1, x_2^* + dx_2) = 0 \quad (2)$$

Poiché il vincolo impone $g(x_1^*, x_2^*) = 0$, possiamo ragionevolmente assumere

$$g(x_1^* + dx_1, x_2^* + dx_2) = 0 \quad (2)$$

L'espansione in serie di Taylor della (2) in un intorno del punto (x_1^*, x_2^*) conduce a

$$g(x_1^* + dx_1, x_2^* + dx_2) \approx g(x_1^*, x_2^*) + \frac{\partial g(x_1^*, x_2^*)}{\partial x_1} dx_1 + \frac{\partial g(x_1^*, x_2^*)}{\partial x_2} dx_2 = 0 \quad (3)$$

Dalla (3) si ottiene:

$$dx_1 = \frac{\frac{\partial g(x_1^*, x_2^*)}{\partial x_2}}{\frac{\partial g(x_1^*, x_2^*)}{\partial x_1}} dx_2 \quad (4)$$

Sostituendo la (4) nella (1) otteniamo:

$$\left(-\frac{\partial f(x_1^*, x_2^*)}{\partial x_1} \frac{\frac{\partial g(x_1^*, x_2^*)}{\partial x_2}}{\frac{\partial g(x_1^*, x_2^*)}{\partial x_1}} + \frac{\partial f(x_1^*, x_2^*)}{\partial x_2} \right) dx_2 = 0 \quad (5)$$

Sostituendo la (4) nella (1) otteniamo:

$$\left(-\frac{\partial f(x_1^*, x_2^*)}{\partial x_1} \frac{\frac{\partial g(x_1^*, x_2^*)}{\partial x_2}}{\frac{\partial g(x_1^*, x_2^*)}{\partial x_1}} + \frac{\partial f(x_1^*, x_2^*)}{\partial x_2} \right) dx_2 = 0 \quad (5)$$

ovvero:

$$\boxed{-\frac{\partial f(x_1^*, x_2^*)}{\partial x_1} \frac{\partial g(x_1^*, x_2^*)}{\partial x_2} + \frac{\partial g(x_1^*, x_2^*)}{\partial x_1} \frac{\partial f(x_1^*, x_2^*)}{\partial x_2} = 0} \quad (6)$$

equazione la cui soluzione conduce all'ottimo.

Ci si propone di estendere le considerazioni svolte per il caso bidimensionale a quelle relative al caso generale:

Si calcoli il vettore \mathbf{x}^ che minimizza $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $f \in C^1$ e soddisfi i vincoli algebrici*

$$g_i(\mathbf{x}^*) = 0 \quad \text{per } i = 1, \dots, m < n. \quad (7)$$

Generalizzando la (3) otteniamo:

$$\begin{aligned} \frac{\partial g_1(\mathbf{x}^*)}{\partial x_1} dx_1 + \frac{\partial g_1(\mathbf{x}^*)}{\partial x_2} dx_2 + \dots + \frac{\partial g_1(\mathbf{x}^*)}{\partial x_n} dx_n &= 0 \\ &\vdots \\ \frac{\partial g_m(\mathbf{x}^*)}{\partial x_1} dx_1 + \frac{\partial g_m(\mathbf{x}^*)}{\partial x_2} dx_2 + \dots + \frac{\partial g_m(\mathbf{x}^*)}{\partial x_n} dx_n &= 0 \end{aligned} \quad (8)$$

METODO VARIAZIONI VINCOLATE - CASO GENERALE

Introdotta il vettore delle incognite

$$d\mathbf{x}_{1:m} = \{ dx_1 \quad dx_2 \quad \dots \quad dx_m \}^T ,$$

la matrice dei coefficienti

$$\mathbf{J} \left(\begin{array}{c} g_1, \dots, g_m \\ x_1, \dots, x_m \end{array} \right) = \left[\begin{array}{cccc} \frac{\partial g_1(\mathbf{x}^*)}{\partial x_1} & \frac{\partial g_1(\mathbf{x}^*)}{\partial x_2} & \dots & \frac{\partial g_1(\mathbf{x}^*)}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m(\mathbf{x}^*)}{\partial x_1} & \frac{\partial g_m(\mathbf{x}^*)}{\partial x_2} & \dots & \frac{\partial g_m(\mathbf{x}^*)}{\partial x_m} \end{array} \right] \quad (9)$$

e il vettore dei termini noti

$$\mathbf{v} = \left\{ \begin{array}{cccc} -\frac{\partial g_1(\mathbf{x}^*)}{\partial x_{m+1}} & -\frac{\partial g_1(\mathbf{x}^*)}{\partial x_{m+2}} & \dots & -\frac{\partial g_1(\mathbf{x}^*)}{\partial x_n} \\ & \vdots & & \\ -\frac{\partial g_m(\mathbf{x}^*)}{\partial x_{m+1}} & -\frac{\partial g_m(\mathbf{x}^*)}{\partial x_{m+2}} & \dots & -\frac{\partial g_m(\mathbf{x}^*)}{\partial x_n} \end{array} \right\} = \left\{ \begin{array}{c} v_1 \\ \vdots \\ v_m \end{array} \right\} \quad (10)$$

e il vettore dei termini noti

$$\mathbf{v} = \left\{ \begin{array}{cccc} -\frac{\partial g_1(\mathbf{x}^*)}{\partial x_{m+1}} & -\frac{\partial g_1(\mathbf{x}^*)}{\partial x_{m+2}} & \cdots & -\frac{\partial g_1(\mathbf{x}^*)}{\partial x_n} \\ & \vdots & & \\ -\frac{\partial g_m(\mathbf{x}^*)}{\partial x_{m+1}} & -\frac{\partial g_m(\mathbf{x}^*)}{\partial x_{m+2}} & \cdots & -\frac{\partial g_m(\mathbf{x}^*)}{\partial x_n} \end{array} \right\} = \left\{ \begin{array}{c} v_1 \\ \vdots \\ v_m \end{array} \right\} \quad (10)$$

il sistema può quindi scriversi nella forma compatta:

$$\mathbf{J} \begin{pmatrix} g_1, \dots, g_m \\ x_1, \dots, x_m \end{pmatrix} d\mathbf{x}_{1 \div m} = \mathbf{v} . \quad (11)$$

METODO VARIAZIONE VINCOLATE - CASO GENERALE

Affinché il sistema (11) ammetta soluzione è necessario sia:

$$\det \mathbf{J} \left(\frac{g_1, \dots, g_m}{x_1, \dots, x_m} \right) \neq 0 \quad (12)$$

METODO VARIAZIONE VINCOLATE - CASO GENERALE

Affinché il sistema (11) ammetta soluzione è necessario sia:

$$\det \mathbf{J} \left(\frac{g_1, \dots, g_m}{x_1, \dots, x_m} \right) \neq 0 \quad (12)$$

La k^{ma} componente contenuta nel vettore $d\mathbf{x}_{1 \div m}$ può esprimersi

$$dx_k = \frac{\det \begin{bmatrix} \frac{\partial g_1(\mathbf{x}^*)}{\partial x_1} & \frac{\partial g_1(\mathbf{x}^*)}{\partial x_2} & \dots & v_1 & \dots & \frac{\partial g_1(\mathbf{x}^*)}{\partial x_m} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial g_m(\mathbf{x}^*)}{\partial x_1} & \frac{\partial g_m(\mathbf{x}^*)}{\partial x_2} & \dots & v_m & \dots & \frac{\partial g_m(\mathbf{x}^*)}{\partial x_m} \end{bmatrix}}{\det \mathbf{J} \left(\frac{g_1, \dots, g_m}{x_1, \dots, x_m} \right)}$$

k
 \downarrow

(13)

METODO VARIAZIONI VINCOLATE - CASO GENERALE

Sviluppando la (13) si ottiene

$$\begin{aligned} & \det \mathbf{J} \left(\frac{g_1, \dots, g_m}{x_1, \dots, x_m} \right) dx_k \\ &= (-1)^{(k+1)+1} \underbrace{\left(\sum_{i=m+1}^n \frac{\partial g_1}{\partial x_i} dx_i \right)}_{-v_1} \det \mathbf{J} \left(\frac{\bigcup_{i=1}^m g_i - g_1}{\bigcup_{i=1}^m x_i - x_k} \right) \\ &+ \dots (-1)^{(k+m)+1} \underbrace{\left(\sum_{i=m+1}^n \frac{\partial g_m}{\partial x_i} dx_i \right)}_{-v_m} \det \mathbf{J} \left(\frac{\bigcup_{i=1}^m g_i - g_m}{\bigcup_{i=1}^m x_i - x_k} \right) \end{aligned}$$

$$\begin{aligned}
&= dx_{m+1} \left((-1)^{(k+1)+1} \frac{\partial g_1}{\partial x_{m+1}} \det \mathbf{J} \left(\frac{\bigcup_{i=1}^m g_i - g_1}{\bigcup_{i=1}^m x_i - x_k} \right) + \dots \right. \\
&+ (-1)^{(k+m)+1} \frac{\partial g_m}{\partial x_{m+1}} \det \mathbf{J} \left(\frac{\bigcup_{i=1}^m g_i - g_m}{\bigcup_{i=1}^m x_i - x_k} \right) \left. + \dots + dx_n(\dots) \right) \\
&= dx_{m+1} \left(\sum_{j=1}^m (-1)^{(k+j)+1} \frac{\partial g_j}{\partial x_{m+1}} \det \mathbf{J} \left(\frac{\bigcup_{i=1}^m g_i - g_j}{\bigcup_{i=1}^m x_i - x_k} \right) + \dots + dx_n(\dots) \right)
\end{aligned}$$

(14)

Sostituendo nella (1) i risultati della (14) si ha:

METODO VARIAZIONI VINCOLATE - CASO GENERALE

Sostituendo nella (1) i risultati della (14) si ha:

$$\begin{aligned}
 & \frac{\partial f}{\partial x_1} \det \mathbf{J} \left(\frac{g_1, \dots, g_m}{x_1, \dots, x_m} \right) dx_1 + \frac{\partial f}{\partial x_2} \det \mathbf{J} \left(\frac{g_1, \dots, g_m}{x_1, \dots, x_m} \right) dx_2 \quad (15) \\
 & + \dots + \frac{\partial f}{\partial x_n} \det \mathbf{J} \left(\frac{g_1, \dots, g_m}{x_1, \dots, x_m} \right) dx_n \\
 & = dx_{m+1} \left(\left(\sum_{i=1}^m \frac{\partial f}{\partial x_i} \sum_{j=1}^m (-1)^{i+j+1} \frac{\partial g_j}{\partial x_{m+1}} \det \mathbf{J} \left(\begin{array}{c} m \\ \bigcup_{t=1}^m g_t - g_j \\ m \\ \bigcup_{t=1}^m x_t - x_i \end{array} \right) \right) \right. \\
 & \left. + \frac{\partial f}{\partial x_{m+1}} \det \mathbf{J} \left(\frac{g_1, \dots, g_m}{x_1, \dots, x_m} \right) \right) + \dots + dx_n(\dots)
 \end{aligned}$$

$$\begin{aligned}
&= \sum_{\ell=m+1}^n dx_{\ell} \left(\left(\sum_{i=1}^m (-1)^i \frac{\partial f}{\partial x_i} \sum_{j=1}^m (-1)^{j+1} \frac{\partial g_j}{\partial x_{\ell}} \det \mathbf{J} \left(\frac{\bigcup_{t=1}^m g_t - g_j}{\bigcup_{t=1}^m x_t - x_i} \right) \right) \right. \\
&\quad \left. + \frac{\partial f}{\partial x_{\ell}} \det \mathbf{J} \left(\frac{g_1, \dots, g_m}{x_1, \dots, x_m} \right) \right) = 0 \tag{16}
\end{aligned}$$

Affinché la (16) sia vera è **necessario** che ogni coefficiente di dx_ℓ , $\forall \ell \in [m + 1 ; n]$ sia nullo. Tali coefficienti possono essere espressi sotto forma di determinante ovvero:

METODO VARIAZIONI VINCOLATE - CASO GENERALE

Affinché la (16) sia vera è **necessario** che ogni coefficiente di dx_ℓ , $\forall \ell \in [m + 1 ; n]$ sia nullo. Tali coefficienti possono essere espressi sotto forma di determinante ovvero:

$$\sum_{\ell=m+1}^n \det \left(\begin{array}{c|ccc} \frac{\partial f}{\partial x_\ell} & \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_m} \\ \hline \frac{\partial g_1}{\partial x_\ell} & \mathbf{J} \left(\frac{g_1, \dots, g_m}{x_1, \dots, x_m} \right) & & \end{array} \right) dx_\ell = 0. \quad (17)$$

Riassumendo si può osservare che:

- Ciascuna equazione di vincolo $g_j(\mathbf{x}) = 0$, per $j = 1, 2, \dots, m$ da luogo ad un'equazione lineare nelle variazioni dx_i , per $i = 1, 2, \dots, n$.

Riassumendo si può osservare che:

- Ciascuna equazione di vincolo $g_j(\mathbf{x}) = 0$, per $j = 1, 2, \dots, m$ da luogo ad un'equazione lineare nelle variazioni dx_i , per $i = 1, 2, \dots, n$.
- Si forma quindi il sistema (11) di m equazioni in $n > m$ incognite. In tal modo i differenziali dx_i presenti df sono espressi in funzione di $n - m$ differenziali indipendenti (dx_{m+1}, \dots, dx_n).

Riassumendo si può osservare che:

- Ciascuna equazione di vincolo $g_j(\mathbf{x}) = 0$, per $j = 1, 2, \dots, m$ da luogo ad un'equazione lineare nelle variazioni dx_i , per $i = 1, 2, \dots, n$.
- Si forma quindi il sistema (11) di m equazioni in $n > m$ incognite. In tal modo i differenziali dx_i presenti df sono espressi in funzione di $n - m$ differenziali indipendenti (dx_{m+1}, \dots, dx_n).
- Imponendo siano nulli i coefficienti dei suddetti $n - m$ differenziali indipendenti si ottengono le **condizioni necessarie** per l'esistenza della soluzione del problema di ottimo:

$$\det \mathbf{J} \left(\begin{array}{c} g_1, \dots, g_m \\ x_1, \dots, x_m \end{array} \right) \neq 0 \quad (18a)$$

$$\det \left(\begin{array}{c|ccc} \frac{\partial f}{\partial x_\ell} & \frac{\partial f}{\partial x_1} & \dots & \frac{\partial f}{\partial x_m} \\ \hline \frac{\partial g_1}{\partial x_\ell} & \mathbf{J} \left(\begin{array}{c} g_1, \dots, g_m \\ x_1, \dots, x_m \end{array} \right) & & \\ \vdots & & & \\ \frac{\partial g_m}{\partial x_\ell} & & & \end{array} \right) = 0 \quad \forall \ell \in [m+1; n] \quad (18b)$$

In teoria, con m vincoli si potrebbe esprimere f solamente in funzione di x_{m+1}, \dots, x_n . L'espansione in serie di Taylor della funzione f può quindi essere scritta come

In teoria, con m vincoli si potrebbe esprimere f solamente in funzione di x_{m+1}, \dots, x_n . L'espansione in serie di Taylor della funzione f può quindi essere scritta come

$$f(\mathbf{x}^* + d\mathbf{x}^*) \approx f(\mathbf{x}^*) + \sum_{i=m+1}^n \left(\frac{\partial f(\mathbf{x}^*)}{\partial x_i} \right) dx_i + \frac{1}{2!} \sum_{i=m+1}^n \sum_{j=m+1}^n \left(\frac{\partial^2 f(\mathbf{x}^*)}{\partial x_i \partial x_j} \right) dx_i dx_j \quad (19)$$

Si osserva che \mathbf{x}^* è un punto stazionario la (19) quindi può essere riscritta come

Si osserva che \mathbf{x}^* è un punto stazionario la (19) quindi può essere riscritta come

$$f(\mathbf{x}^* + d\mathbf{x}^*) - f(\mathbf{x}^*) \approx \frac{1}{2!} \sum_{i=m+1}^n \sum_{j=m+1}^n \left(\frac{\partial^2 f(\mathbf{x}^*)}{\partial x_i \partial x_j} \right) dx_i dx_j = Q. \quad (20)$$

Si osserva che \mathbf{x}^* è un punto stazionario la (19) quindi può essere riscritta come

$$f(\mathbf{x}^* + d\mathbf{x}^*) - f(\mathbf{x}^*) \approx \frac{1}{2!} \sum_{i=m+1}^n \sum_{j=m+1}^n \left(\frac{\partial^2 f(\mathbf{x}^*)}{\partial x_i \partial x_j} \right) dx_i dx_j = Q. \quad (20)$$

Condizione sufficiente Q sia un minimo (massimo) è che Q sia definita positiva (negativa).

Le caratteristiche del metodo dei moltiplicatori di Lagrange saranno inizialmente discusse per il caso di ottimizzazione di una funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f \in C^1$ soggetta ad un vincolo $g : \mathbb{R}^2 \rightarrow \mathbb{R}$, $g \in C^1$

Le caratteristiche del metodo dei moltiplicatori di Lagrange saranno inizialmente discusse per il caso di ottimizzazione di una funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f \in C^1$ soggetta ad un vincolo $g : \mathbb{R}^2 \rightarrow \mathbb{R}$, $g \in C^1$

Caso bidimensionale ($n = 2$, $m = 1$)

Minimizzare $f(x_1, x_2)$

soggetta a

$$g(x_1, x_2) = 0$$

Definiamo, secondo Lagrange, la nuova funzione obiettivo L :

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2) \quad (21)$$

con λ quantità incognita che prende il nome di moltiplicatore di Lagrange.

Definiamo, secondo Lagrange, la nuova funzione obiettivo L :

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2) \quad (21)$$

con λ quantità incognita che prende il nome di moltiplicatore di Lagrange.

Osserviamo che:

- il punto di ottimo (x_1^*, x_2^*) della funzione L coincide con quello della funzione f ;

Definiamo, secondo Lagrange, la nuova funzione obiettivo L :

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2) \quad (21)$$

con λ quantità incognita che prende il nome di moltiplicatore di Lagrange.

Osserviamo che:

- il punto di ottimo (x_1^*, x_2^*) della funzione L coincide con quello della funzione f ;
- l'artificio introdotto da Lagrange permette di trasformare il problema di ottimo vincolato in uno di ottimizzazione non vincolata.

Applicando alla (21) le condizioni necessarie di estremo si ottiene il sistema:

MOLTIPLICATORI DI LAGRANGE

Applicando alla (21) le condizioni necessarie di estremo si ottiene il sistema:

$$\frac{\partial L(x_1, x_2, \lambda)}{\partial x_1} = \frac{\partial f(x_1, x_2)}{\partial x_1} + \lambda \frac{\partial g(x_1, x_2)}{\partial x_1} = 0 \quad (22a)$$

$$\frac{\partial L(x_1, x_2, \lambda)}{\partial x_2} = \frac{\partial f(x_1, x_2)}{\partial x_2} + \lambda \frac{\partial g(x_1, x_2)}{\partial x_2} = 0 \quad (22b)$$

$$\frac{\partial L(x_1, x_2, \lambda)}{\partial \lambda} = g(x_1, x_2) = 0 \quad (22c)$$

MOLTIPLICATORI DI LAGRANGE

Applicando alla (21) le condizioni necessarie di estremo si ottiene il sistema:

$$\frac{\partial L(x_1, x_2, \lambda)}{\partial x_1} = \frac{\partial f(x_1, x_2)}{\partial x_1} + \lambda \frac{\partial g(x_1, x_2)}{\partial x_1} = 0 \quad (22a)$$

$$\frac{\partial L(x_1, x_2, \lambda)}{\partial x_2} = \frac{\partial f(x_1, x_2)}{\partial x_2} + \lambda \frac{\partial g(x_1, x_2)}{\partial x_2} = 0 \quad (22b)$$

$$\frac{\partial L(x_1, x_2, \lambda)}{\partial \lambda} = g(x_1, x_2) = 0 \quad (22c)$$

Eliminando λ dalle (22a) e (22b) si ottiene la (6)

$$-\frac{\partial f(x_1^*, x_2^*)}{\partial x_1} \frac{\partial g(x_1^*, x_2^*)}{\partial x_2} + \frac{\partial g(x_1^*, x_2^*)}{\partial x_1} \frac{\partial f(x_1^*, x_2^*)}{\partial x_2} = 0$$

MOLTIPLICATORI DI LAGRANGE - CASO GENERALE

Estendiamo al caso generale di n variabili ed m vincoli le considerazioni svolte nel precedentemente ($m < n$):

MOLTIPLICATORI DI LAGRANGE - CASO GENERALE

Estendiamo al caso generale di n variabili ed m vincoli le considerazioni svolte nel precedentemente ($m < n$):

Minimizzare $f(\mathbf{x})$

soggetta a

$$g_j(\mathbf{x}) = 0, \quad (j = 1, 2, \dots, m)$$

La nuova funzione obiettivo L costituita da $(n + m)$ variabili è definita come segue:

MOLTIPLICATORI DI LAGRANGE - CASO GENERALE

Estendiamo al caso generale di n variabili ed m vincoli le considerazioni svolte nel precedentemente ($m < n$):

Minimizzare $f(\mathbf{x})$

soggetta a

$$g_j(\mathbf{x}) = 0, \quad (j = 1, 2, \dots, m)$$

La nuova funzione obiettivo L costituita da $(n + m)$ variabili è definita come segue:

$$L(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_m) = f(\mathbf{x}) + \lambda_1 g_1(\mathbf{x}) + \dots + \lambda_m g_m(\mathbf{x}) \quad (23)$$

Le condizioni necessarie affinché L abbia un estremo sono le seguenti:

$$\frac{\partial L}{\partial x_i} = \frac{\partial f(\mathbf{x})}{\partial x_i} + \sum_{j=1}^m \lambda_j \frac{\partial g_j(\mathbf{x})}{\partial x_i} = 0, \quad \text{per } i = 1, 2, \dots, n \quad (24a)$$

$$\frac{\partial L}{\partial \lambda_j} = g_j(\mathbf{x}) = 0, \quad \text{per } j = 1, 2, \dots, m \quad (24b)$$

Le condizioni necessarie affinché L abbia un estremo sono le seguenti:

$$\frac{\partial L}{\partial x_i} = \frac{\partial f(\mathbf{x})}{\partial x_i} + \sum_{j=1}^m \lambda_j \frac{\partial g_j(\mathbf{x})}{\partial x_i} = 0, \quad \text{per } i = 1, 2, \dots, n \quad (24a)$$

$$\frac{\partial L}{\partial \lambda_j} = g_j(\mathbf{x}) = 0, \quad \text{per } j = 1, 2, \dots, m \quad (24b)$$

Le equazioni (24) formano un sistema di $(n + m)$ equazioni nelle $(n + m)$ incognite x_1, x_2, \dots, x_n e $\lambda_1, \lambda_2, \dots, \lambda_m$

Teorema

La condizione sufficiente affinché la funzione $f(\mathbf{x})$ abbia un minimo (massimo) relativo in \mathbf{x}^* è che la funzione quadratica

$$Q = \sum_{i=1}^n \sum_{j=1}^m \frac{\partial^2 L}{\partial x_i \partial x_j} dx_i dx_j \quad (25)$$

valutata in $\mathbf{x} = \mathbf{x}^*$ sia definita positiva (negativa), per tutti i valori di $d\mathbf{x}$ che soddisfano i vincoli.

In particolare affinché Q sia definita positiva (negativa) per tutte le variazioni possibili $d\mathbf{x}$, è necessario che siano positive (negative) tutte le radici del polinomio in μ ottenuto espandendo il determinante

$$\det \begin{bmatrix} L_{11} - \mu & L_{12} & L_{13} & \dots & L_{1n} & g_{11} & g_{21} & \dots & g_{m1} \\ L_{21} & L_{22} - \mu & L_{23} & \dots & L_{2n} & g_{12} & g_{22} & \dots & g_{m2} \\ \vdots & & & & & & & & \\ L_{n1} & L_{n2} & L_{n3} & \dots & L_{nn} - \mu & g_{1n} & g_{2n} & \dots & g_{mn} \\ g_{11} & g_{12} & g_{13} & \dots & g_{1n} & 0 & 0 & \dots & 0 \\ g_{21} & g_{22} & g_{23} & \dots & g_{2n} & 0 & 0 & \dots & 0 \\ \vdots & & & & & & & & \\ g_{m1} & g_{m2} & g_{m3} & \dots & g_{mn} & 0 & 0 & \dots & 0 \end{bmatrix} = 0 \quad (26)$$

$$\text{con } L_{ij} = \frac{\partial^2 L(\mathbf{x}^*, \lambda^*)}{\partial x_i \partial x_j} \text{ e } g_{ij} = \frac{\partial g_i}{\partial x_j}.$$

PROGRAMMAZIONE NON LINEARE

La formulazione generale dei problemi di programmazione non lineare (NLP) si presenta nella forma:

$$\begin{aligned} &\text{Minimizzare } \theta(\mathbf{x}) \text{ soggetta ai seguenti vincoli} \\ &g_i(\mathbf{x}) \leq 0 \quad i \in M = \{1, 2, \dots, m\} \\ &h_j(\mathbf{x}) = 0 \quad j \in K = \{1, 2, \dots, k\} \end{aligned} \tag{27}$$

dove $\theta(\mathbf{x})$, $g_i(\mathbf{x})$ e $h_j(\mathbf{x})$ sono funzioni definite su uno spazio Euclideo n -dimensionale E^n ed hanno derivate parziali prime continue su E^n .

I criteri di ottimo, possono essere suddivisi in due categorie:

- Necessari: Se \mathbf{x}^* è soluzione del problema (27), le relazioni che legano $\theta(\mathbf{x}^*)$, $g_j(\mathbf{x}^*)$ e $h_j(\mathbf{x}^*)$ sono considerate condizioni necessarie.
- Sufficienti: In taluni problemi sarà possibile stabilire un insieme di condizioni grazie alle quali si garantisce che un punto \mathbf{x}^* sia un estremo relativo.

Teorema

Siano \mathbf{A} , \mathbf{B} e \mathbf{C} matrici reali e costanti con \mathbf{A} non identicamente nulla. O il sistema:

$$\mathbf{y}^T \mathbf{A} < 0 \quad \mathbf{y}^T \mathbf{B} \leq 0 \quad \mathbf{y}^T \mathbf{C} = 0 \quad (28)$$

ha soluzione $\bar{\mathbf{y}}$, o il sistema

$$\mathbf{A}\mathbf{z}_1 + \mathbf{B}\mathbf{z}_2 + \mathbf{C}\mathbf{z}_3 = 0 \quad \mathbf{z}_1 \geq 0 \quad \mathbf{z}_1 \neq 0 \quad \mathbf{z}_2 \geq 0 \quad (29)$$

ha soluzione $\bar{\mathbf{z}}_1, \bar{\mathbf{z}}_2$ e $\bar{\mathbf{z}}_3$. Entrambe le possibilità non possono sussistere.

LEMMA I - PARTE I

Siano $f_i(\mathbf{x})$ $i \in L = \{1, 2, \dots, \ell\} \neq \emptyset$ e $h_j, j = \{1, 2, \dots, k\}$, funzioni definite in un insieme aperto D di E^n ed aventi derivate parziali continue in D .

Si consideri il sistema

$$f_i(\mathbf{x}) = 0 \quad i \in L \quad (30a)$$

$$h_j(\mathbf{x}) = 0 \quad j \in K \quad (30b)$$

con soluzione $\mathbf{x} \in D$ e sia

$$f_i(\mathbf{x}) < 0 \quad i \in L \quad (31a)$$

$$h_j(\mathbf{x}) = 0 \quad j \in K \quad (31b)$$

un sistema che non ha soluzione in D .

Allora il sistema

$$\mathbf{y}^T \nabla f_i(\mathbf{x}^*) < 0 \quad i \in L \quad (32a)$$

$$\mathbf{y}^T \nabla h_j(\mathbf{x}^*) < 0 \quad j \in K \quad (32b)$$

non avrà soluzione in $\mathbf{y} \in E^n$, assunto che

$$\nabla h_j(\mathbf{x}^*), j \in K \quad (33)$$

siano linearmente indipendenti.

LEMMA II

Siano valide le ipotesi del Lemma I. In tal caso esisteranno dei vettori $\bar{\mathbf{r}} \in E^\ell$, $\bar{\mathbf{s}} \in E^k$, tali che

$$\sum_{i=1}^{\ell} \bar{r}_i \nabla f_i(\mathbf{x}^*) + \sum_{j=1}^k \bar{s}_j \nabla h_j(\mathbf{x}^*) = 0 \quad (34a)$$

$$\bar{\mathbf{r}} \geq 0 \quad (34b)$$

$$\begin{bmatrix} \bar{\mathbf{r}} \\ \bar{\mathbf{s}} \end{bmatrix} \neq 0 \text{ .}^1 \quad (34c)$$

¹Con la presente notazione si intende che non tutte le componenti del vettore possono annullarsi.

CONDIZIONI GENERALIZZATE NECESSARIE DI FRITZ JOHN

Se \mathbf{x}^* è soluzione del problema (27), allora esistono dei vettori

$$\bar{\mathbf{u}} = \begin{Bmatrix} \bar{u}_0 \\ \bar{u}_1 \\ \vdots \\ \bar{u}_m \end{Bmatrix}, \quad \bar{\mathbf{u}} \in E^{m+1} \quad \bar{\mathbf{v}} = \begin{Bmatrix} \bar{v}_1 \\ \bar{v}_2 \\ \vdots \\ \bar{v}_k \end{Bmatrix}, \quad \bar{\mathbf{v}} \in E^k \quad (35)$$

tali che

$$\begin{aligned} \bar{u}_0 \nabla \theta(\mathbf{x}^*) + \sum_{i=1}^m \bar{u}_i \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^k \bar{v}_j \nabla h_j(\mathbf{x}^*) &= 0 \\ \sum_{i=1}^m \bar{u}_i g_i(\mathbf{x}^*) &= 0 \\ \bar{\mathbf{u}} &\geq 0 \\ \begin{bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{v}} \end{bmatrix} &\neq 0 \end{aligned}$$

Dimostrazione

Sia

$$\bar{M} = \{i \mid i \in M, g_i(\mathbf{x}^*) = 0\} \quad (36)$$

e

$$D = \left\{ \mathbf{x}^* \mid \mathbf{x}^* \in E^n, g_i(\mathbf{x}^*) < 0, i \in M - \bar{M} \right\} \quad (37)$$

Dimostrazione

Sia

$$\bar{M} = \{i \mid i \in M, g_i(\mathbf{x}^*) = 0\} \quad (36)$$

e

$$D = \left\{ \mathbf{x}^* \mid \mathbf{x}^* \in E^n, g_i(\mathbf{x}^*) < 0, i \in M - \bar{M} \right\} \quad (37)$$

Poniamo inoltre:

$$f_1(\mathbf{x}) = \theta(\mathbf{x}) - \theta(\mathbf{x}^*) \quad (38a)$$

$$f_i(\mathbf{x}) = g_{P_i}(\mathbf{x}), \quad P_i \in \bar{M} \quad (i = 2, 3, \dots, \ell) \quad (38b)$$

Considerata la definizione di $f_1(\mathbf{x})$, l'insieme $L = \{1, 2, \dots, \ell\}$ non è vuoto (indipendentemente dal fatto che \overline{M} sia vuoto o meno).

Il sistema

$$f_i(\mathbf{x}) = 0, \quad i \in L \quad (39a)$$

$$h_j(\mathbf{x}) = 0, \quad j \in K \quad (39b)$$

ha soluzione $\mathbf{x}^* \in D$. Per contro, il sistema

$$f_i(\mathbf{x}) < 0, \quad i \in L \quad (40a)$$

$$h_j(\mathbf{x}) = 0, \quad j \in K \quad (40b)$$

NON ha soluzione in D .

Considerata la definizione di $f_1(\mathbf{x})$, l'insieme $L = \{1, 2, \dots, \ell\}$ non è vuoto (indipendentemente dal fatto che \overline{M} sia vuoto o meno).

Il sistema

$$f_i(\mathbf{x}) = 0, \quad i \in L \quad (39a)$$

$$h_j(\mathbf{x}) = 0, \quad j \in K \quad (39b)$$

ha soluzione $\mathbf{x}^* \in D$. Per contro, il sistema

$$f_i(\mathbf{x}) < 0, \quad i \in L \quad (40a)$$

$$h_j(\mathbf{x}) = 0, \quad j \in K \quad (40b)$$

NON ha soluzione in D . Infatti, se il sistema delle (40) avesse soluzione, \mathbf{x}^* non sarebbe un minimo e quindi le ipotesi finora fatte non sarebbero valide.

Abbiamo appena verificato che le ipotesi del Lemma II sono verificate, quindi possiamo affermare l'esistenza di due vettori

$$\bar{\mathbf{r}} = \begin{bmatrix} \bar{r}_0 \\ \bar{r}_{P_2} \\ \bar{r}_{P_3} \\ \vdots \\ \bar{r}_{P_\ell} \end{bmatrix} \quad (\bar{\mathbf{r}} \in E^\ell) \quad (41)$$

e

$$\bar{\mathbf{s}} = \begin{bmatrix} \bar{s}_1 \\ \bar{s}_2 \\ \vdots \\ \bar{s}_k \end{bmatrix} \quad (\bar{\mathbf{s}} \in E^k) \quad (42)$$

Abbiamo appena verificato che le ipotesi del Lemma II sono verificate, quindi possiamo affermare l'esistenza di due vettori

$$\bar{\mathbf{r}} = \begin{bmatrix} \bar{r}_0 \\ \bar{r}_{P_2} \\ \bar{r}_{P_3} \\ \vdots \\ \bar{r}_{P_\ell} \end{bmatrix} \quad (\bar{\mathbf{r}} \in E^\ell) \quad (41)$$

e

$$\bar{\mathbf{s}} = \begin{bmatrix} \bar{s}_1 \\ \bar{s}_2 \\ \vdots \\ \bar{s}_k \end{bmatrix} \quad (\bar{\mathbf{s}} \in E^k) \quad (42)$$

tali che

$$\bar{r}_0 \nabla \theta(\mathbf{x}^*) + \sum_{i=2}^{\ell} \bar{r}_{P_i} \nabla g_{P_i}(\mathbf{x}^*) + \sum_{j=1}^k \bar{s}_j \nabla h_j(\mathbf{x}^*) = 0 \quad (43)$$

con $\bar{\mathbf{r}} \geq 0$ e $[\bar{\mathbf{u}} \quad \bar{\mathbf{v}}]^T \neq 0$.

Si tenga presente che $P_i \in \overline{M}$ e in generale $\nabla g_{P_i}(\mathbf{x}^*) \neq 0$.

Si tenga presente che $P_i \in \overline{M}$ e in generale $\nabla g_{P_i}(\mathbf{x}^*) \neq 0$.
Se definiamo $\bar{\mathbf{u}} \in E^{m+1}$ e $\bar{\mathbf{v}} \in E^k$ nella maniera seguente

$$\bar{u}_0 = \bar{r}_0 \quad (44)$$

$$\bar{u}_i = \begin{cases} 0 & \text{per } i \in M - \overline{M} \\ \bar{r}_i & \text{per } i \in \overline{M} \end{cases} \quad (45)$$

(l'esistenza degli \bar{r}_i è garantita dal Lemma 2)

$$\bar{\mathbf{v}} = \bar{\mathbf{s}} \quad (46)$$

si avrà che:

1. $\sum_{i=1}^m \bar{u}_i g_i(\mathbf{x}^*) = 0$ per la definizione di $\bar{\mathbf{u}}$, in quanto per $i \in M - \bar{M}$, $u_i = 0$;
2. è possibile riscrivere la relazione (43) come segue

$$\bar{u}_0 \nabla \theta(\mathbf{x}^*) + \sum_{i=1}^m \bar{u}_i \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^k \bar{v}_j \nabla h_j(\mathbf{x}^*) = 0 \quad (47)$$

con

$$\bar{\mathbf{u}} \geq 0 \quad (48)$$

$$\begin{bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{v}} \end{bmatrix} \neq 0 \quad (49)$$