Il caso generale: funzioni di n variabili con m vincoli ($m < n$)

Problema

Dato un aperto $X \subseteq \mathbb{R}^n$ e date due funzioni di classe C^1 $f : X \to \mathbb{R}$ e $\boldsymbol{g} : X \to \mathbb{R}^m$, con $m < n$, determinare gli estremi di f, ristretta all'insieme (**vincolo**):

$$E_0 = \{\boldsymbol{x} \in \mathbb{R}^n : g(\boldsymbol{x}) = 0\}$$

Se $\boldsymbol{g} = (g_1, g_2, \ldots, g_m)$, E_0 è l'insieme dei punti di $\mathbb{R}^n$ che soddisfano le m equazioni:

$$\begin{cases} g_1(x_1, x_2, \ldots, x_n) = 0 \\ g_2(x_1, x_2, \ldots, x_n) = 0 \\ \vdots \\ g_m(x_1, x_2, \ldots, x_n) = 0 \end{cases}$$

Il caso $n = 2$, $m = 1$ è quello precedente.

Se dal sistema precedente riusciamo a ricavare esplicitamente m variabili in termini delle altre $n - m$, riduciamo il problema di ottimizzazione vincolata ad un problema di ottimizzazione ancora vincolata, ma di dimensione inferiore.

Di solito si utilizza il metodo dei *moltiplicatori di Lagrange*, che estende quello trattato precedentemente.

Definizione

$\boldsymbol{x}^0 = (x_1^0, x_2^0, \dots, x_n^0) \in X$ un punto **regolare** di E_0 se la matrice Jacobiana di $\boldsymbol{g}$:

$$\mathbf{Dg} = \begin{pmatrix} D_{x_1} g_1 & D_{x_2} g_1 & \dots & D_{x_n} g_1 \\ D_{x_1} g_2 & D_{x_2} g_2 & \dots & D_{x_n} g_2 \\ \vdots & \vdots & & \vdots \\ D_{x_1} g_m & D_{x_2} g_m & \dots & D_{x_n} g_m \end{pmatrix}$$

ha caratteristica m in $\boldsymbol{x}^0$.

Affinché la matrice $\mathbf{Dg}$ abbia caratteristica m in $\boldsymbol{x}^0$, i vettori $\nabla g_j(\boldsymbol{x}^0)$, per $j = 1, \ldots, m$ devono essere linearmente indipendenti. Esiste quindi un minore di ordine m, estratto da $\mathbf{Dg}(\boldsymbol{x}^0)$, diverso da zero. Si può supporre che questo minore sia $\frac{\partial(g_1, g_2, \ldots, g_m)}{\partial(x_1, x_2, \ldots, x_m)}$, cioè quello individuato dalle prime m colonne di $\mathbf{Dg}$.

Teorema del Dini (o della funzione implicita) in due dimensioni: stabilisce quando il luogo di zeri di un'equazione implicita si può esplicitare rispetto ad una variabile.
Sia $g : X \subseteq \mathbb{R}^2 \to \mathbb{R}$, X aperto. Supponiamo che:

1. g e g_y siano continue in X;
2. nel punto $(x_0, y_0) \in X$ si abbia $g(x_0, y_0) = 0$ e $g_y(x_0, y_0) \neq 0$.

Allora esistono un intorno U di x_0 e un'unica funzione $f : U \to \mathbb{R}$, continua in U, tale che $y_0 = f(x_0)$ e che $g(x, f(x)) = 0$, $\forall x \in U$. Se inoltre g_x è continua in X (quindi $g \in C^1(X)$), allora $f \in C^1(X)$ e vale:

$$f'(x) = -\frac{g_x(x, f(x))}{g_y(x, f(x))}, \quad \forall x \in U$$

Nel nostro caso grazie al teorema del Dini:
il sistema definisce $x_1, \ldots, x_m$, in funzione di $x_{m+1}, \ldots, x_n$ in un intorno di $\boldsymbol{x}^0$:

$$\begin{gathered} x_1 = h_1(x_{m+1}, \ldots, x_n) \\ x_2 = h_2(x_{m+1}, \ldots, x_n) \\ \vdots \\ x_m = h_m(x_{m+1}, \ldots, x_n) \end{gathered}$$

in modo tale che $x_j^0 = h_j(x_{m+1}^0, \ldots, x_n^0)$, $\forall j = 1, \ldots, m$.
Si dice che: in un intorno di $\boldsymbol{x}^0$, E_0 è una **varietà di dimensione** $n - m$.
Quindi per individuare un punto su E_0 in un intorno di $\boldsymbol{x}^0$ occorrono $n - m$ parametri.

La varietà di dimensione $n-1$ si chiama *ipersuperficie* in $\mathbb{R}^n$.

Le varietà di dimensione 1 si chiamano *curve* in $\mathbb{R}^n$.

In generale una curva in $\mathbb{R}^n$ è assegnata da un vettore:

$$\boldsymbol{r}(t) = (x_1(t), x_2(t), \ldots, x_n(t)), \qquad \text{con } t \in I \subseteq \mathbb{R}$$

La curva si dice regolare se $\boldsymbol{r} \in C^1(I)$ e se il vettore tangente o velocità:

$$\boldsymbol{r}'(t) = (x_1'(t), x_2'(t), \ldots, x_n'(t)) \neq \mathbf{0}, \quad \forall t \in I$$

Definizione

Il punto regolare $\boldsymbol{x}^0 \in X$ si dice **critico** o **stazionario** per f, vincolato alla varietà $E_{\mathbf{0}}$, se per ogni curva regolare $\boldsymbol{r} = \boldsymbol{r}(t)$ tale che:

1. $\boldsymbol{r}$ è definita in un intorno di I_0 di $t = 0$ e $\boldsymbol{r}(0) = \boldsymbol{x}^0$;
2. la curva è contenuta in $E_{\mathbf{0}}$ e cioè
   $g_j(\boldsymbol{r}(t)) = g_j(x_1(t), \ldots, x_n(t)) = 0, \ \forall j = 1, \ldots, m, \ \forall t \in I_0$

posto $\phi(t) = f(\boldsymbol{r}(t))$, si ha $\phi'(0) = 0$.

Teorema

Se $\boldsymbol{x}^0 \in X$ è punto di estremo per f, vincolato alla varietà $E_{\mathbf{0}}$, ed è punto regolare per $E_{\mathbf{0}}$, allora $\boldsymbol{x}^0$ è punto critico vincolato (a $E_{\mathbf{0}}$).

Dimostrazione:
Se $\boldsymbol{x}^0$ è di massimo (o di minimo) vincolato e $\boldsymbol{r} = \boldsymbol{r}(t)$ è una curva soddisfacente le condizioni 1) e 2), allora $\phi(t) = f(\boldsymbol{r}(t))$ ha un massimo (o un minimo) libero in $t = 0$ e perciò $\phi'(0) = 0$. $\square$

Teorema

Sia $\boldsymbol{x}^0 \in X$ un punto regolare per $E_{\mathbf{0}}$. Allora $\boldsymbol{x}^0$ è un punto critico vincolato a $E_{\mathbf{0}}$ per f se e solo se esistono m numeri reali $\lambda_1^0, \lambda_2^0, \ldots, \lambda_m^0$, detti **moltiplicatori di Lagrange**, tali che:

$$\nabla f(\boldsymbol{x}^0) = \sum_{j=1}^{m} \lambda_j^0 \nabla g_j(\boldsymbol{x}^0)$$

Le direzioni corrispondenti ai vettori $\boldsymbol{v} \in \mathbb{R}^n$ tali che $\langle \nabla g_j(\boldsymbol{x}^0), \boldsymbol{v} \rangle = 0, \ \forall j = 1, \ldots, m$ si dicono *direzioni tangenziali ai vincoli.*

Un punto regolare $\boldsymbol{x}^0 \in X$ è stazionario se e solo se la derivata di f lungo ogni direzione tangenziale ai vincoli è nulla.

Definiamo la Lagrangiana:

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = \mathcal{L}(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m) := f(\boldsymbol{x}) - \sum_{j=1}^{m} \lambda_j g_j(\boldsymbol{x})$$

il Teorema afferma che se $\boldsymbol{x}^0$ è regolare, allora è critico condizionato se e solo se esiste $\boldsymbol{\lambda}^0 = (\lambda_1^0, \ldots, \lambda_m^0)$ tale che $(\boldsymbol{x}^0, \boldsymbol{\lambda}^0) \in \mathbb{R}^n \times \mathbb{R}^m$ sia punto critico libero per $\mathcal{L}$.
I punti critici liberi per $\mathcal{L}$ sono soluzioni del sistema di $n+m$ equazioni in $n+m$ incognite:

$$\begin{cases} D_{x_1}\mathcal{L} = D_{x_1} f - \sum_{j=1}^{m} \lambda_j D_{x_1} g_j = 0 \\ D_{x_2}\mathcal{L} = D_{x_2} f - \sum_{j=1}^{m} \lambda_j D_{x_2} g_j = 0 \\ \vdots \\ D_{x_n}\mathcal{L} = D_{x_n} f - \sum_{j=1}^{m} \lambda_j D_{x_n} g_j = 0 \\ D_{\lambda_1}\mathcal{L} = -g_1 = 0 \\ D_{\lambda_2}\mathcal{L} = -g_2 = 0 \\ \vdots \\ D_{\lambda_m}\mathcal{L} = -g_m = 0 \end{cases}$$

Condizioni sufficienti

Come stabilire la natura del punto critico?
Analizziamo il caso di un solo vincolo ($m = 1$).

Sia $\boldsymbol{x}^0 \in \mathbb{R}^n$ un punto critico per f condizionato al vincolo $g(x) = 0$ e sia λ_0 il corrispondente moltiplicatore di Lagrange.
Vale il seguente:

Teorema

Siano $f, g : X \to \mathbb{R}$, X aperto di $\mathbb{R}^n$, di classe C^2.
Se la forma quadratica:

$$\sum_{i,j=1}^{n} \left[f_{x_i x_j}(\boldsymbol{x}^0) - \lambda^0 g_{x_i x_j}(\boldsymbol{x}^0)\right] h_i h_j = \langle \left[\mathbf{H}_f(\boldsymbol{x}^0) - \lambda^0 \mathbf{H}_g(\boldsymbol{x}^0)\right] \mathbf{h}, \mathbf{h} \rangle$$

ristretta all'insieme dei vettori $\mathbf{h} \in \mathbb{R}^n$ tangenziali al vincolo in $\boldsymbol{x}^0$, (cioè $\langle \nabla g(\boldsymbol{x}^0), \mathbf{h} \rangle = 0$), è definita negativa (positiva), allora $\boldsymbol{x}^0$ è punto di massimo (minimo) locale forte vincolato.

Lemma

Sia $q(\boldsymbol{h}) = \langle \mathbf{Ah}, \mathbf{h} \rangle$ una forma quadratica in $\mathbb{R}^n$ e sia $\mathbf{b} \in \mathbb{R}^n$, $\mathbf{b} = (b_1, b_2, \ldots, b_n)$, $b_1 \neq 0$. Allora la f.q. $q(\mathbf{h})$, soggetta al vincolo lineare $\langle \mathbf{b}, \mathbf{h} \rangle = 0$, è definita positiva se sono **negativi** tutti i minori principali di nord-ovest, di ordine maggiore di 2, della seguente matrice:

$$\begin{pmatrix} 0 & b_1 & \ldots & b_n \\ b_1 & & & \\ \vdots & & \mathbf{A} & \\ b_n & & & \end{pmatrix}$$

È definita negativa se i suddetti minori si susseguono a segni alterni a partire dal primo (di ordine 3) **positivo**.

Il lemma si applica nel nostro caso con $\mathbf{A} = \mathbf{H}_f(\boldsymbol{x}^0) - \lambda^0 \mathbf{H}_g(\boldsymbol{x}^0)$ e $\mathbf{b} = \nabla g(\boldsymbol{x}^0)$; poiché $\nabla g(\boldsymbol{x}^0) \neq \mathbf{0}$, si può sempre supporre che $D_{x_1} g(\boldsymbol{x}^0) \neq 0$.

Esempio:

Massimizzare $V(x,y,z) = xyz$ nell'ottante positivo di $\mathbb{R}^3$, sotto la condizione $g(x,y,z) = xy + xz + yz - A = 0$ ($A > 0$).
Poiché $\nabla g = (y+z, x+z, z+y)$ si annulla solo in $(0,0,0)$ sulla superficie non vi sono punti singolari.
Cerchiamo i punti critici. Il sistema è assegnato dalle equazioni:

$$\begin{cases} V_x - \lambda g_x = yz - \lambda(y+z) = 0 \\ V_y - \lambda g_y = xz - \lambda(x+z) = 0 \\ V_z - \lambda g_z = xy - \lambda(x+y) = 0 \end{cases}$$

alle quali va aggiunta $xy + xz + yz - A = 0$. Troviamo che l'unico punto $x = y = z = \sqrt{\frac{A}{2}}$ con $\lambda = \sqrt{\frac{A}{8}}$. Dobbiamo stabilire la natura del punto critico $\left(\sqrt{\frac{A}{2}}, \sqrt{\frac{A}{2}}, \sqrt{\frac{A}{2}}\right)$.

Abbiamo che:

$$\begin{cases} V_{xx} - \lambda g_{xx} = V_{yy} - \lambda g_{yy} = V_{zz} - \lambda g_{zz} = 0 \\ V_{xy} - \lambda g_{xy} = z - \lambda \\ V_{xz} - \lambda g_{xz} = y - \lambda \\ V_{yz} - \lambda g_{yz} = x - \lambda \\ \nabla g = (y + z, x + z, z + y) \end{cases}$$

Sostituendo $\lambda = \sqrt{\frac{A}{8}}$, $\left(\sqrt{\frac{A}{2}}, \sqrt{\frac{A}{2}}, \sqrt{\frac{A}{2}}\right)$, otteniamo:

$$\begin{pmatrix} 0 & \sqrt{2A} & \sqrt{2A} & \sqrt{2A} \\ \sqrt{2A} & 0 & \sqrt{\frac{A}{8}} & \sqrt{\frac{A}{8}} \\ \sqrt{2A} & \sqrt{\frac{A}{8}} & 0 & \sqrt{\frac{A}{8}} \\ \sqrt{2A} & \sqrt{\frac{A}{8}} & \sqrt{\frac{A}{8}} & 0 \end{pmatrix} = \sqrt{2A} \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

Il determinante di questa matrice è negativo, il suo minore principale di nord-ovest di ordine 3 è positivo, quindi la f.q. risulta definita negativa e quindi $\boldsymbol{x}^0$ è punto di massimo locale forte vincolato.

Vincoli di disuguaglianza

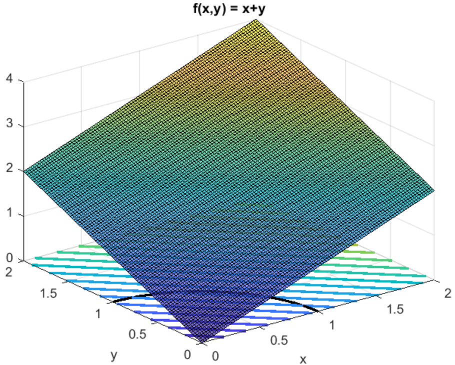
Nell'analisi economica ci si imbatte frequentemente nel problema di ottimizzare una funzione $f(\boldsymbol{x})$ quando la variabile $\boldsymbol{x}$ è soggetta ad uno o più vincoli del tipo $\varphi(\boldsymbol{x}) \leq 0$, oltre alla condizione di non-negatività delle sue componenti, cioè $x_j \geq 0$, $j = 1, 2, \ldots, n$. Questo è un tipico problema di programmazione matematica.

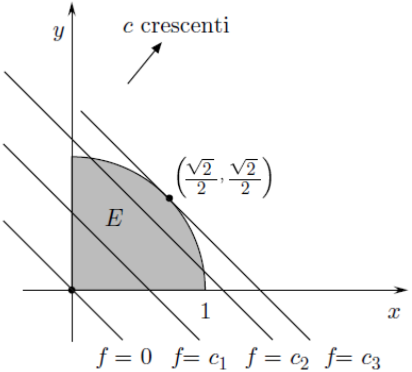
Usando la teoria sviluppata finora un problema di questo tipo ci obbligherebbe ad un procedimento abbastanza lungo:

1. la ricerca degli estremi (liberi) di f nell'interno della regione individuata dai vincoli dove f è differenziabile;
2. la ricerca degli estremi di f sulla frontiera di tale regione;
3. la considerazione dei punti in cui f non è differenziabile.

Se la funzione obiettivo ha un'espressione analitica semplice si può ricorrere a considerazioni grafiche.

Esempio:
Determinare gli estremi di $f(x, y) = x + y$ nella regione
$E = \{(x, y) \in \mathbb{R}^2 \, : \, x^2 + y^2 - 1 \leq 0, \, x \geq 0, \, y \geq 0\}$.
Per individuare i punti di estremo è sufficiente determinare la linea di massimo e di minimo livello che interseca la regione E.





Tale metodo però è legato a una situazione troppo particolare, quindi occorrono dei metodi adattabili a casi di sufficiente generalità.

Formulazione generale del problema

Siano $f, \varphi_1, \varphi_2, \ldots, \varphi_m : \mathbb{R}^n \to \mathbb{R}$ funzioni differenziabili.

$(\mathcal{P})$: Massimizzare f ristretta alla regione:

$$E = \{\boldsymbol{x} \in \mathbb{R}^n : \varphi_j(\boldsymbol{x}) \leq 0, j = 1, \ldots, m\}$$

E viene detta regione *ammissibile*.

Attenzione:

- Non c'è alcuna relazione tra il numero n di variabili ed il numero m di vincoli. Supporremo sempre $E \neq \emptyset$;
- per semplicità di trattazione assumiamo la differenziabilità della funzione obiettivo e delle funzioni φ_j in tutto $\mathbb{R}^n$;
- tutti i risultati si estendono facilmente al problema della minimizzazione di f in E in quanto $\min_E f = -\max_E(-f)$. Gli estremi si intendono globali rispetto alla regione E;
- scrivere i vincoli nella forma $\varphi_j \leq 0$ anziché $\varphi_j \geq 0$ è del tutto convenzionale e generale. Appare ovvio che ogni vincolo può essere espresso nella forma $\varphi_j \leq 0$. In ogni caso osserviamo che se $\varphi_j(\boldsymbol{x}^0) = 0$ e $\nabla\varphi_j(\boldsymbol{x}^0) \neq \mathbf{0}$, $\nabla\varphi_j(\boldsymbol{x}^0)$ è normale all'insieme di livello $\{\boldsymbol{x} \in \mathbb{R}^n : \varphi_j(\boldsymbol{x}) = 0\}$ e diretto verso *l'esterno* della regione $E_j = \{\boldsymbol{x} \in \mathbb{R}^n, \varphi_j(\boldsymbol{x}) \leq 0\}$, con $E = \cap_{j=1}^m E_j$.

Definizione

Sia $\boldsymbol{x}^0$ una soluzione di $(\mathcal{P})$.

Il vincolo è **attivo** in $\boldsymbol{x}^0$ se il punto soddisfa almeno una delle disuguaglianze $\varphi_j(\boldsymbol{x}) \leq 0$ (o $\varphi_j(\boldsymbol{x}) \geq 0$) con l'uguaglianza, ossia se vale $\varphi_j(\boldsymbol{x}^0) = 0$ per almeno un j.

Se la soluzione $\boldsymbol{x}^0$ è tale che $\varphi_j(\boldsymbol{x}^0) < 0$ (o $\varphi_j(\boldsymbol{x}^0) > 0$), $\forall j = 1, \ldots, m$, allora il vincolo è **inattivo** in $\boldsymbol{x}^0$.

Il vincolo è attivo se la funzione obiettivo raggiunge il suo valore massimo (o minimo) su un punto di frontiera del vincolo, mentre è inattivo se il punto estremo è interno.

Indichiamo con $J(\boldsymbol{x})$ è l'insieme degli indici $j \in \{1, 2, \ldots, m\}$ per cui i vincoli corrispondenti sono *attivi* in $\boldsymbol{x}$.

Lemma

Se per ogni $j \in J(\boldsymbol{x}^0)$ esiste $\overline{\boldsymbol{x}}$ tale che:

- o $\varphi_j(\overline{\boldsymbol{x}}) < 0$ e φ_j è convessa,
- oppure $\varphi_j(\overline{\boldsymbol{x}}) \leq 0$ e φ_j è affine,

allora i vincoli sono qualificati in $\boldsymbol{x}^0$.

Lemma

Se i vettori $\nabla\varphi_j(\boldsymbol{x}^0)$ per $j \in J(\boldsymbol{x}^0)$ sono linearmente indipendenti, allora i vincoli sono qualificati in $\boldsymbol{x}^0$.

Teorema (di Karush-Kuhn-Tucker)

Sia $\boldsymbol{x}^0$ soluzione di $(\mathcal{P})$.
Se i vincoli sono qualificati in $\boldsymbol{x}^0$, allora per ogni $j \in J(\boldsymbol{x}^0)$ esiste $\lambda_j^0 \geq 0$ tale che:

$$\nabla f(\boldsymbol{x}^0) = \sum_{j \in J(\boldsymbol{x}^0)} \lambda_j^0 \nabla \varphi_j(\boldsymbol{x}^0)$$

La formula del teorema fa intervenire solo i vincoli attivi in $\boldsymbol{x}^0$.
Per il calcolo, poiché $\boldsymbol{x}^0$ non è conosciuto a priori, dobbiamo introdurre la Lagrangiana:

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) := f(\boldsymbol{x}) - \sum_{j=1}^{m} \lambda_j \varphi_j(\boldsymbol{x})$$

nella quale intervengono tutti i vincoli.

Attenzione:
La condizione necessaria del teorema equivale a richiedere che $\boldsymbol{x}^0$ e $\boldsymbol{\lambda}^0$ siano soluzione del seguente sistema (che esprime le *condizioni di Karush-Kuhn-Tucker*):

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x_k} = \frac{\partial f}{\partial x_k} - \sum_{j=1}^{m} \lambda_j \frac{\partial \varphi_j}{\partial x_k} = 0 \quad k = 1, \ldots, n \\ \frac{\partial \mathcal{L}}{\partial \lambda_j} = -\varphi_j \geq 0 \quad j = 1, \ldots, m \\ \lambda_j \geq 0; \quad j = 1, \ldots, m \\ \lambda_j \varphi_j(\boldsymbol{x}) = 0 \quad j = 1, \ldots, m \end{cases}$$

Se $j \notin J(\boldsymbol{x}^0)$, dalle ultime due condizioni segue che $\lambda_j = 0$. Questo comporta che nelle prime due condizioni compaiano soltanto i vincoli attivi.

Esempio: Minimizzare $g(x_1, x_2) = (x_1 - 1)^3 + x_2 + 1$ nella regione $E = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0, x_2 \geq 0\}$.
Usiamo la relazione $\min_E g = -\max_E(-g)$. In questo modo il problema diventa:

$$
\begin{aligned}
&f(x_1, x_2) = -g(x_1, x_2) = -(x_1 - 1)^3 - x_2 - 1, \\
&\varphi_1(x_1, x_2) = -x_1, \quad \varphi_2(x_1, x_2) = -x_2.
\end{aligned}
$$

Notiamo che i vincoli sono affini, quindi qualificati. La $\mathcal{L}(x_1, x_2, \lambda_1, \lambda_2) = -(x_1 - 1)^3 - x_2 - 1 + \lambda_1 x_1 + \lambda_2 x_2$ e quindi

$$
\begin{cases}
\frac{\partial \mathcal{L}}{\partial x_1} = -3(x_1 - 1)^2 + \lambda_1 = 0 \\
\frac{\partial \mathcal{L}}{\partial x_2} = -1 + \lambda_2 = 0 \\
x_1 \geq 0, \quad x_2 \geq 0 \\
\lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_1 x_1 = 0, \lambda_2 x_2 = 0.
\end{cases}
$$

da cui si ricava $\lambda_2 = 1, x_2 = 0$. Se $x_1 = 0, \lambda_1 = 3$. Se $\lambda_1 = 0, x_1 = 1$. Per le condizioni di Karush-Kuhn-Tucker abbiamo che i possibili punti sono $(0, 0)$ e $(1, 0)$, con $f(0, 0) = 0$ e $f(1, 0) = -1$. Osserviamo che $f_{|E} \leq 0$ e pertanto 0 è massimo globale per f ovvero minimo globale per g.

Osservazione:
Le condizioni espresse dal sistema si scrivono per la stessa maniera per i problemi seguenti:
$\max f$ con i vincoli nella forma $\varphi_j \leq 0$
e
$\min f$ con i vincoli nella forma $\varphi_j \geq 0$.

Osservazione:
I numeri λ_j che compaiono nella formula di Karush-Kuhn-Tucker prendono il nome di *moltiplicatori di Karush-Kuhn-Tucker*.

Teorema

Supponiamo che f sia concava e che le funzioni φ_j, $j = 1, \ldots, m$, siano convesse. Sia $\boldsymbol{x}^0 \in E$ un punto in cui i vincoli sono qualificati. Allora $\boldsymbol{x}^0$ è soluzione di $(\mathcal{P})$ se e solo se, per ogni $j \in J(\boldsymbol{x}^0)$, esiste $\lambda_j^0 \geq 0$ tale che valga:

$$\nabla f(\boldsymbol{x}^0) = \sum_{i \in J(\boldsymbol{x}^0)} \lambda_j^0 \nabla \varphi_j(\boldsymbol{x}^0)$$

Dimostrazione:
"$\Leftarrow$": Se f è concava, $\forall \boldsymbol{x} \in E$ e dalla formula sopra:

$$f(\boldsymbol{x}) \leq f(\boldsymbol{x}^0) + \langle \nabla f(\boldsymbol{x}^0), \boldsymbol{x} - \boldsymbol{x}^0 \rangle = f(\boldsymbol{x}^0) + \sum_{j \in J(\boldsymbol{x}^0)} \lambda_j^0 \langle \nabla \varphi_j(\boldsymbol{x}^0), \boldsymbol{x} - \boldsymbol{x}^0 \rangle$$

Siccome φ_j sono convesse e siccome $\varphi(\boldsymbol{x}^0) = 0$ se $j \in J(\boldsymbol{x}^0)$, abbiamo che: $\varphi_j(\boldsymbol{x}) \geq \langle \nabla \varphi_j(\boldsymbol{x}^0), \boldsymbol{x} - \boldsymbol{x}^0 \rangle$.
Inoltre essendo $\lambda_j^0 \geq 0$ e $\varphi_j(\boldsymbol{x}) \leq 0$ per $x \in E$, abbiamo che $f(\boldsymbol{x}) \leq f(\boldsymbol{x}^0) + \sum_{j \in J(\boldsymbol{x}^0)} \lambda_j^0 \varphi_j(\boldsymbol{x}) \leq f(\boldsymbol{x}^0)$.
Quindi $\boldsymbol{x}^0$ è soluzione di $(\mathcal{P})$. $\square$

Esempio:
Massimizzare $f(x_1, x_2) = x_1 + \log(1 + x_2)$ nell'insieme
$E = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 - 1 \leq 0, x_1 \geq 0, x_2 \geq 0\}$.

f è concava essendo somma di due funzioni concave.
Inoltre $\varphi_1 = x_1 + x_2 - 1$, $\varphi_2 = -x_1$, $\varphi_3 = -x_2$ sono affini e quindi convesse.
La Lagrangiana è:

$$\mathcal{L}(x_1, x_2, \lambda_1, \lambda_2, \lambda_3) = x_1 + \log(1 + x_2) - \lambda_1(x_1 + x_2 - 1) + \lambda_2 x_1 + \lambda_3 x_2$$

Le condizioni di Karush-Kuhn-Tucker (necessarie e sufficienti) sono:

$$\begin{cases} 1 - \lambda_1 + \lambda_2 = 0 \\ \frac{1}{1+x_2} - \lambda_1 + \lambda_3 = 0 \\ \lambda_1 \geq 0,\ \lambda_2 \geq 0,\ \lambda_3 \geq 0 \\ \lambda_1(x_1 + x_2 - 1) = 0,\ \lambda_2 x_1 = 0,\ \lambda_3 x_2 = 0 \end{cases}$$

L'unica soluzione è $(1, 0)$ con i moltiplicatori $(1, 0, 0)$. Il valore $f(1, 0) = 1$ è quindi di massimo globale per f.