VI. CONVEXIDAD Y EXTREMOS DE FUNCIONES DIFERENCIABLES

6.1 Funciones Convexas

Definición 6.1.1: Una función f definida en una parte convexa (ver definición 4.2.1) Q de un e.v. $\vec{E}$ con valores en $I\!R$, se dirá convexa en Q si para todo $\vec{x}, \vec{y} \in Q$ se tiene la desigualdad

$$f(\lambda\vec{x} + (1-\lambda)\vec{y}) \leq \lambda f(\vec{x}) + (1-\lambda)f(\vec{y}) \quad \text{para todo} \quad \lambda \in [0,1] \qquad 6.1.1$$

Si la desigualdad anterior es estricta cuando $\vec{x} \neq \vec{y}$ y $\lambda \in]0,1[$, diremos que la función f es estrictamente convexa.

Definición 6.1.2: Se llama epigrafo de una función f definida en una parte A de un e.v. $\vec{E}$ con valores en $I\!R$, al conjunto

$$epi(f) = \{(\vec{x}, z) \in A \times I\!R : f(\vec{x}) \leq z\} \qquad 6.1.2$$

Geométricamente, esto se interpreta como el conjunto de los puntos $(\vec{x}, z) \in A \times I\!R$ que están "sobre" el grafo de f (recordemos que el grafo de una función f es el conjunto $\{(\vec{x}, z) \in A \times I\!R : f(\vec{x}) = z\}$)

Nota 6.1.1: De acuerdo a la definición anterior, la desigualdad 6.1.1 es equivalente a

$$(\lambda\vec{x} + (1-\lambda)\vec{y}\ ,\ \lambda f(\vec{x}) + (1-\lambda)f(\vec{y})) \in epi(f)$$

que escrito de otro modo toma la forma

$$\lambda(\vec{x}, f(\vec{x})) + (1-\lambda)(\vec{y}, f(\vec{y})) \in epi(f).$$

Vemos entonces que f será convexa si y solo si para todo $\vec{x}, \vec{y} \in Q$ el trazo que une los puntos $(\vec{x}, f(\vec{x}))$ e $(\vec{y}, f(\vec{y}))$ en el grafo de f, pertenece al epigrafo de f. No es difícil entonces ver que la convexidad de una función f es equivalente a la convexidad del conjunto $epi(f)$ en $\vec{E} \times I\!R$.

Teorema 6.1.1: Dadas dos funciones convexas (resp. estrictamente convexas) f, g definidas en un conjunto convexo Q de un e.v. $\vec{E}$ con valores en $I\!R$ y dado $t \in I\!R_+$, entonces

i) la función $f+g$ es convexa (resp. estrictamente convexa).
ii) la función tf es convexa (resp. estrictamente convexa).

Demostración: Dados $\vec{x}, \vec{y} \in Q$ y $\lambda \in [0,1]$, se tiene

i)
$$\begin{aligned}[f+g](\lambda\vec{x} + (1-\lambda)\vec{y}) &= f(\lambda\vec{x} + (1-\lambda)\vec{y}) + g(\lambda\vec{x} + (1-\lambda)\vec{y}) \\ &\leq \lambda f(\vec{x}) + (1-\lambda)f(\vec{y}) + \lambda g(\vec{x}) + (1-\lambda)g(\vec{y}) \\ &= \lambda[f+g](\vec{x}) + (1-\lambda)[f+g](\vec{y})\end{aligned}$$

ii)
$$\begin{aligned}[tf](\lambda\vec{x} + (1-\lambda)\vec{y}) &= t(f(\lambda\vec{x} + (1-\lambda)\vec{y})) \\ &\leq t(\lambda f(\vec{x}) + (1-\lambda)f(\vec{y})) \\ &= \lambda[tf](\vec{x}) + (1-\lambda)[tf](\vec{y}).\end{aligned}$$

Teorema 6.1.2: Dada una familia $\{g_t\}_{t\in T}$ de funciones convexas definidas en un conjunto convexo Q de un e.v. $\vec{E}$ con valores en $I\!R$, si para cada $\vec{x} \in Q$ el conjunto $\{g_t(\vec{x}) : t \in T\}$ es acotado superiormente, entonces la función $f : Q \to I\!R$ definida por

$$f(\vec{x}) = \sup\{g_t(\vec{x}) : t \in T\}$$

es convexa.

Demostración: De acuerdo con la nota 6.1.1 basta con demostrar que $epi(f)$ es un conjunto convexo en $\vec{E} \times I\!R$. Vamos a mostrar entonces que $epi(f) = \bigcap_{t\in T} epi(g_t)$ y, como la intersección de conjuntos convexos es siempre convexa, habremos concluido que $epi(f)$ es convexo. La igualdad en cuestion es una consecuencia inmediata de la cadena de implicaciones

$$\begin{aligned}(\vec{x}, z) \in epi(f) \Leftrightarrow f(\vec{x}) \leq z \Leftrightarrow g_t(\vec{x}) \leq z \quad \forall t \in T \Leftrightarrow (\vec{x}, z) \in epi(g_t) \quad \forall t \in T\\ \Leftrightarrow (\vec{x}, z) \in \bigcap_{t\in T} epi(g_t)\end{aligned}$$

6.2 Caracterización de funciones convexas diferenciables

Nota 6.2.1: En los dos teoremas que siguen daremos una caracterización importante de la convexidad (resp. estricta convexidad) para funciones diferenciables.

Teorema 6.2.1: Sea f una función diferenciable definida en un abierto convexo Q de un e.v.n. $\vec{E}$ con valores en $I\!R$. Una condición necesaria y suficiente para que f sea convexa (resp. estrictamente convexa) es que

$$f(\vec{x}) \geq f(\vec{y}) + Df(\vec{y})(\vec{x} - \vec{y}) \quad \text{para todo} \quad \vec{x}, \vec{y} \in Q \tag{6.2.1}$$

$$(\text{resp. } f(\vec{x}) > f(\vec{y}) + Df(\vec{y})(\vec{x} - \vec{y}) \quad \text{para todo} \quad \vec{x}, \vec{y} \in Q \text{ con } \vec{x} \neq \vec{y})$$

Demostración: i) Supongamos primero que f es convexa. Entonces para todo $\vec{x}, \vec{y} \in Q$ y todo $\lambda \in]0, 1]$ se tiene de acuerdo a la desigualdad 6.1.1.

$$f(\vec{y} + \lambda(\vec{x} - \vec{y})) - f(\vec{y}) \leq \lambda[f(\vec{x}) - f(\vec{y})]$$

Dividiendo por λ y tomando límite cuando $\lambda \to 0$ con $\lambda > 0$, obtenemos la desigualdad $Df(\vec{y}; \vec{x} - \vec{y}) \leq f(\vec{x}) - f(\vec{y})$, que de acuerdo a la fórmula 5.2.5 es equivalente a 6.2.1 .

ii) Supongamos ahora que f verifica 6.2.1. Sean $\vec{x}, \vec{y} \in Q$ y $\vec{z} := \vec{y} + \lambda(\vec{x} - \vec{y})$ con $\lambda \in [0, 1]$. Entonces

$$\begin{aligned}f(\vec{x}) \geq f(\vec{z}) + Df(\vec{z})(\vec{x} - \vec{z})\\ f(\vec{y}) \geq f(\vec{z}) + Df(\vec{z})(\vec{y} - \vec{z})\end{aligned}$$

y multiplicando estas desigualdades por λ y $(1 - \lambda)$ respectivamente y sumándola, se obtiene gracias a la linealidad de la función $Df(\vec{z})$

$$\begin{aligned}\lambda f(\vec{x}) + (1 - \lambda)f(\vec{y}) &\geq f(\vec{z}) + Df(\vec{z})(\lambda(\vec{x} - \vec{z}) + (1 - \lambda)(\vec{y} - \vec{z}))\\ &= f(\vec{z}) + Df(\vec{z})(\vec{0})\\ &= f(\lambda\vec{x} + (1 - \lambda)\vec{y})\end{aligned}$$

que corresponde exactamente a la convexidad de f.

Nota 6.2.2: Si en el teorema anterior $\vec{E} = I\!R^n$, la fórmula 6.2.1 se escribe

$$f(\vec{x}) \geq f(\vec{y}) + \langle \nabla f(\vec{y}), \vec{x} - \vec{y} \rangle \quad \text{para todo} \quad \vec{x}, \vec{y} \in Q \qquad 6.2.2$$

$$(\text{ resp. } f(\vec{x}) > f(\vec{y}) + \langle \nabla f(\vec{y}), \vec{x} - \vec{y} \rangle \quad \text{para todo} \quad \vec{x}, \vec{y} \in Q \text{ con } \vec{x} \neq \vec{y})$$

Nota 6.2.3: En la nota 5.2.2 veíamos que la ecuación $z = h(\vec{x})$, donde $h : \vec{E} \to \vec{F}$ es la función definida por $h(\vec{x}) := f(\vec{y}) + Df(\vec{y})(\vec{x} - \vec{y})$ (con $\vec{y} \in Q$ fijo), representa el hiperplano afín tangente al grafo de una función diferenciable $f : Q \to I\!R$ en $(\vec{y}, f(\vec{y}))$. Esto nos muestra, de acuerdo al teorema anterior, que una función diferenciable f es convexa si y solo si para todo $\vec{y} \in Q$ el hiperplano afín tangente al grafo de f en $(\vec{y}, f(\vec{y}))$ está por debajo del grafo de f, es decir, para todo $\vec{y} \in Q$ la función lineal afín h minora a la función f.

Teorema 6.2.2: Sea f una función diferenciable definida en un abierto convexo Q de un e.v.n. $\vec{E}$ con valores en $I\!R$. Una condición necesaria y suficiente para que f sea convexa (resp. estrictamente convexa) es que

$$[Df(\vec{x}) - Df(\vec{y})](\vec{x} - \vec{y}) \geq 0 \quad \text{para todo } \vec{x}, \vec{y} \in Q \qquad 6.2.3$$

$$(\text{ resp. } [Df(\vec{x}) - Df(\vec{y})](\vec{x} - \vec{y}) > 0 \quad \text{para todo} \quad \vec{x}, \vec{y} \in Q \quad \text{con} \quad \vec{x} \neq \vec{y})$$

Demostración: i) Supongamos que f es convexa. Aplicando dos veces el teorema anterior podemos escribir para todo $\vec{x}, \vec{y} \in Q$

$$f(\vec{x}) - Df(\vec{y})(\vec{x} - \vec{y}) \geq f(\vec{y})$$
$$f(\vec{y}) + Df(\vec{x})(\vec{x} - \vec{y}) \geq f(\vec{x})$$

sumando y teniendo en cuenta la linealidad de $Df(\vec{x})$ y $Df(\vec{y})$ se obtiene la desigualdad 6.2.3 .
ii) Supongamos ahora que f verifica 6.2.3. Dados $\vec{x}, \vec{y} \in Q$ definamos la función $\phi : [0,1] \to I\!R$ por $\phi(\lambda) := f(\vec{x} + \lambda(\vec{y} - \vec{x}))$. Del teorema 5.5.1 vemos que ϕ es derivable y $\phi'(\lambda) = Df(\vec{x} + \lambda(\vec{y} - \vec{x}))(\vec{y} - \vec{x})$. Aplicando entonces 6.2.3 deducimos que para todo $\lambda \in [0,1]$ se tiene

$$\begin{aligned} \phi'(\lambda) - \phi'(0) &= [Df(\vec{x} + \lambda(\vec{y} - \vec{x})) - Df(\vec{x})](\vec{y} - \vec{x}) \\ &= \frac{1}{\lambda}[Df(\vec{x} + \lambda(\vec{y} - \vec{x})) - Df(\vec{x})](\vec{x} + \lambda(\vec{y} - \vec{x}) - \vec{x}) \\ &\geq 0 \end{aligned}$$

Aplicando ahora el teorema del valor medio a la función ϕ en $[0,1]$, vemos que existe $\bar{\lambda} \in]0,1[$ tal que $\phi(1) - \phi(0) = \phi'(\bar{\lambda})$ y de la desigualdad anterior concluimos que

$$\phi(1) - \phi(0) \geq \phi'(0)$$

que corresponde exactamente a la desigualdad 6.2.1. La función f será entonces convexa.

Nota 6.2.4: Si en el teorema anterior $\vec{E} = I\!R^n$, la fórmula 6.2.3 se escribe

$$\langle \nabla f(\vec{x}) - \nabla f(\vec{y}), \vec{x} - \vec{y} \rangle \geq 0 \quad \text{para todo} \quad \vec{x}, \vec{y} \in Q \qquad 6.2.4$$

$$(\text{ resp. } \langle \nabla f(\vec{x}) - \nabla f(\vec{y}), \vec{x} - \vec{y} \rangle > 0 \quad \text{para todo } \vec{x}, \vec{y} \in Q \quad \text{con} \quad \vec{x} \neq \vec{y})$$

En el próximo teorema daremos una caracterización importante de la convexidad (y de la estricta convexidad) para funciones de clase $\mathcal{C}^2$. Como los vectores en $I\!R^n$ los denotamos por una fila, el producto de una matriz H de $n \times n$ con un vector $\vec{\delta}$ de $I\!R^n$ lo escribimos $\vec{\delta}H$ y el producto escalar del vector $\vec{\delta}H$ con un vector $\vec{u}$ en $I\!R^n$ lo escribiremos indistintamente $\vec{\delta}H\vec{u}^t$ o bien $< \vec{\delta}H, \vec{u} >$, donde $\vec{u}^t$ representa el traspuesto del vector $\vec{u}$, es decir el vector $\vec{u}$ escrito como columna.

Teorema 6.2.3: Sea f una función de clase $\mathcal{C}^2$ definida en un abierto convexo Q de $I\!R^n$ con valores en $I\!R$. Una condición necesaria y suficiente para que f sea convexa (resp. estrictamente convexa) es que para todo $\vec{x} \in Q$ su Hessiano $H(\vec{x})$ sea definido no negativo (resp. definido positivo), es decir

$$\vec{\delta}H(\vec{x})\vec{\delta}^t \geq 0 \quad \text{para todo} \quad \vec{\delta} \in I\!R^n \qquad 6.2.5$$

$$(\text{ resp. } \vec{\delta}H(\vec{x})\vec{\delta}^t > 0 \quad \text{para todo} \quad \vec{\delta} \in I\!R^n \quad \text{con} \quad \vec{\delta} \neq \vec{0})$$

Demostración: i) Supongamos que f es convexa. Dados $\vec{x} \in Q$ y $\vec{\delta} \in I\!R^n$ $(\vec{\delta} \neq \vec{0})$, definamos la función $\varphi : [0, \varepsilon] \to I\!R$, donde $\varepsilon > 0$ es tal que $\vec{x} + \varepsilon\vec{\delta} \in Q$, por $\varphi(\lambda) := \langle \nabla f(\vec{x} + \lambda\vec{\delta}), \vec{\delta} \rangle$. Como f es de clase $\mathcal{C}^2$, usando los teoremas 5.9.4 y 5.5.1 se deduce que φ es derivable y $\varphi'(\lambda) = \langle \vec{\delta}H(\vec{x} + \lambda\vec{\delta}), \vec{\delta} \rangle$. La desigualdad 6.2.5 equivale entonces a demostrar que $\varphi'(0) \geq 0$. Puesto que

$$\varphi'(0) = \lim_{\lambda \to 0} \frac{\varphi(\lambda) - \varphi(0)}{\lambda}$$

y como del teorema anterior se tiene

$$\begin{aligned} \varphi(\lambda) - \varphi(0) &= \langle \nabla f(\vec{x} + \lambda\vec{\delta}) - \nabla f(\vec{x}) \ , \ \vec{\delta} \rangle \\ &= \frac{1}{\lambda} \langle \nabla f(\vec{x} + \lambda\vec{\delta}) - \nabla f(\vec{x}), \vec{x} + \lambda\vec{\delta} - \vec{x} \rangle \\ &\geq 0 \end{aligned}$$

concluimos que $\varphi'(0) \geq 0$.
ii) Supongamos ahora que para todo $\vec{x} \in Q$ la matriz $H(\vec{x})$ es definida nonegativa. Dados $\vec{x}, \vec{y} \in Q$, si definimos $\vec{\delta} := \vec{y} - \vec{x}$ y φ como en la parte i) con $\varepsilon = 1$, vemos que aplicando el teorema del valor medio a φ en $[0, 1]$, existirá $\bar{\lambda} \in]0, 1[$ tal que

$$\varphi(1) - \varphi(0) = \varphi'(\bar{\lambda}).$$

Puesto que $\varphi'(\lambda) =< \vec{\delta}H(\vec{x} + \lambda\vec{\delta}), \vec{\delta} >\geq 0$ para todo $\lambda \in]0, 1[$, concluimos que
$\varphi(1) - \varphi(0) \geq 0$ lo que equivale exactamente a la desigualdad 6.2.3 del teorema anterior. La función f será entonces convexa.

Nota 6.2.5: Del teorema anterior vemos que una forma cuadrática q definida en $I\!R^n$ con valores en $I\!R$, esto es $q(\vec{x}) := \alpha + \langle \vec{c}, \vec{x} \rangle + \vec{x}H\vec{x}^t$, es convexa si y solo si la matriz H es definida nonegativa.

Nota 6.2.6: Del teorema anterior y de la nota 5.9.5 vemos que una función f de clase $\mathcal{C}^2$ definida en un abierto convexo Q de $I\!R^n$ con valores en $I\!R$, será convexa si y solo si en todo punto de Q la aproximación cuadrática de f es convexa.

6.3 Funciones Concavas

Definición 6.3.1: Una función f definida en una parte convexa Q de un e.v.$\vec{E}$ con valores en $I\!R$, se dirá cóncava (resp. estrictamente cóncava) en Q si la función $-f$ es convexa (resp. estrictamente convexa) en Q.

Nota 6.3.1: Todo lo dicho en las dos secciones anteriores para las funciones convexas (resp. estrictamente convexas) es válido para las funciones cóncavas (resp. estrictamente cóncavas) cambiando f por $-f$. Así entonces, en el teorema 6.1.1 habrá que cambiar el término convexas por cóncavas; en el teorema 6.1.2 habrá que cambiar el término convexas por cóncavas; en el teorema 6.2.1 habrá que cambiar el término convexa por cóncava y el sentido de la desigualdad 6.2.1; en el teorema 6.2.2 habrá que cambiar el término convexa por cóncava y el sentido de la desigualdad 6.2.3 y finalmente; en el teorema 6.2.3 habrá que cambiar el término convexa por cóncava, $H(\vec{x})$ definido no negativo por $H(\vec{x})$ definido no positivo y el sentido de la desigualdad 6.2.5 .

6.4 Mínimos y máximos de una función

Definición 6.4.1: Dada una función f definida en un conjunto A de un e.v.n. $\vec{E}$ con valores en $I\!R$, un punto $\vec{x}^* \in A$ se dirá mínimo local (resp. mínimo local estricto) de f en una parte D de A si $\vec{x}^* \in D$ y existe $B(\vec{x}^*, \varepsilon)$ tal que

$$f(\vec{x}^*) \leq f(\vec{x}) \quad \text{para todo} \quad \vec{x} \in B(\vec{x}^*, \varepsilon) \cap D \qquad 6.4.1$$

$$(\text{ resp.} \quad f(\vec{x}^*) < f(\vec{x}) \quad \text{para todo} \quad \vec{x} \in B(\vec{x}^*, \varepsilon) \cap D \quad \text{con} \quad \vec{x} \neq \vec{x}^*)$$

Definición 6.4.2: Si la desigualdad 6.4.1 de la definición anterior se verifica en sentido contrario, el punto $\vec{x}^* \in D$ se dirá máximo local (resp. máximo local estricto) de f en D.

Nota 6.4.1: De las definiciones anteriores se deduce fácilmente que un punto $\vec{x}^* \in D$ es máximo local (resp. máximo local estricto) de f en D si y solo si es mínimo local (resp. mínimo local estricto) de la función $-f$.

Definición 6.4.3: Dada una función f definida en un conjunto A de un e.v.n. $\vec{E}$ con valores en $I\!R$, un punto $\vec{x}^* \in A$ se dirá mínimo (resp. mínimo estricto) de f en una parte D de A si $\vec{x}^* \in D$ y

$$f(\vec{x}^*) \leq f(\vec{x}) \quad \text{para todo} \quad \vec{x} \in D \qquad 6.4.2$$

$$(\text{ resp.} \quad f(\vec{x}^*) < f(\vec{x}) \quad \text{para todo} \quad \vec{x} \in D \quad \text{con} \quad \vec{x} \neq \vec{x}^*)$$

Definición 6.4.4: Si la desigualdad 6.4.2 de la definición anterior se verifica en sentido contrario, el punto $\vec{x}^* \in D$ se dirá máximo (resp. máximo estricto) de f en D.

Nota 6.4.2: De las dos últimas definiciones se deduce fácilmente que un punto $\vec{x}^* \in D$ es un máximo (resp. máximo estricto) de f en D si y solo si es un mínimo (resp. mínimo estricto) de la función $-f$.

Nota 6.4.3: De las definiciones anteriores se deduce fácilmente que todo mínimo (resp. mínimo estricto) es también mínimo local (resp. mínimo local estricto) y que todo máximo (resp. máximo estricto) es también máximo local (resp. máximo local estricto).

Teorema 6.4.1: Sea f una función convexa definida en una parte convexa Q de un e.v.n. $\vec{E}$ con valores en $I\!R$. Entonces, todo mínimo local (resp. mínimo local estricto) de f en una parte convexa D de Q será un mínimo (resp. mínimo estricto) de f en D.

Demostración: Sea $\vec{x}^*$ un mínimo local de f en D. Existe entonces $B(\vec{x}^*, \varepsilon)$ tal que

$$f(\vec{x}^*) \leq f(\vec{x}) \quad \text{para todo} \quad \vec{x} \in B(\vec{x}^*, \varepsilon) \cap D \tag{$*$}$$

Para concluir debemos demostrar que la desigualdad anterior también se tiene para todo $\vec{x} \in D \setminus B(x^*, \varepsilon)$. Sea entonces $\vec{x} \in D$ con $\vec{x} \notin B(\vec{x}^*, \varepsilon)$. Como D es un conjunto convexo y $\lambda := \frac{\varepsilon}{\|\vec{x}-\vec{x}^*\|} \in [0,1]$, es fácil ver que $\lambda\vec{x} + (1-\lambda)\vec{x}^* \in D \cap B(\vec{x}^*, \varepsilon)$. Así entonces, de $(*)$ y de la convexidad de f, se tiene

$$f(\vec{x}^*) \leq f(\lambda\vec{x} + (1-\lambda)\vec{x}^*) \leq \lambda f(\vec{x}) + (1-\lambda) f(\vec{x}^*) \leq f(\vec{x})$$

que es lo que queriamos demostrar.

Teorema 6.4.2: Sea f una función diferenciable definida en un abierto A de un e.v.n.$\vec{E}$ con valores en $I\!R$. Una condición necesaria para que $\vec{x}^* \in A$ sea un mínimo local de f en una parte convexa D de A, es que $\vec{x}^* \in D$ y

$$Df(\vec{x}^*)(\vec{x} - \vec{x}^*) \geq 0 \quad \text{para todo } \vec{x} \in D \tag{6.4.3}$$

Demostración: Sea $\vec{x}^*$ un mínimo local de f en D. Dado $\vec{x} \in D$ definamos la función $\phi : [0,1] \to I\!R$ por $\phi(\lambda) := f(\vec{x}^* + \lambda(\vec{x} - \vec{x}^*))$. Del teorema 5.5.1 vemos que ϕ es derivable y que $\phi'(0) = Df(\vec{x}^*)(\vec{x} - \vec{x}^*)$. Por otra parte, como $\phi(\lambda) \geq \phi(0)$ para todo $\lambda \in [0,1]$, concluimos que

$$\phi'(0) = \lim_{\substack{\lambda \to 0 \\ \lambda > 0}} \frac{\phi(\lambda) - \phi(0)}{\lambda} \geq 0$$

lo que corresponde exactamente a la desigualdad 6.4.3.

Nota 6.4.4: De acuerdo a la interpretación que se dió en la nota 5.1.1 a la cantidad $Df(\vec{x}^*; \vec{x} - \vec{x}^*) = Df(\vec{x}^*)(\vec{x} - \vec{x}^*)$ (ver fórmula 5.2.5), vemos que la desigualdad 6.4.3 corresponde al hecho que las pendientes de f en $\vec{x}^*$, en todas las direcciones interiores al convexo D, son no negativas.

Nota 6.4.5: Es fácil ver que la condición 6.4.3 no es suficiente para que $\vec{x}^* \in D$ sea un mínimo local de f en D. Un ejemplo simple que muestra este hecho está dado por la función $f : I\!R \to I\!R$ definida por $f(\vec{x}) = x^3$, que verifica 6.4.3 para $\vec{x}^* = 0$.
El próximo teorema nos muestra que para una clase particular de funciones esta condición es suficiente para que $\vec{x}^* \in D$ sea mínimo.

Nota 6.4.6: Si en el teorema anterior $\vec{E} = I\!R^n$, la desigualdad 6.4.3 se escribe

$$\langle \nabla f(\vec{x}^*), \vec{x} - \vec{x}^* \rangle \geq 0 \quad \text{para todo} \quad \vec{x} \in D \qquad 6.4.4$$

Teorema 6.4.3: Sea f una función convexa y diferenciable definida en un abierto convexo Q de un e.v.n. $\vec{E}$ con valores en $I\!R$. Una condición necesaria y suficiente para que $\vec{x}^* \in Q$ sea un mínimo de f en una parte convexa D en Q, es que $\vec{x}^* \in D$ y verifique 6.4.3.

Demostración: i) Del teorema anterior y, aplicando el teorema 6.4.1, deducimos que la condición es necesaria. ii) supongamos ahora que $\vec{x}^* \in D$ verifica 6.4.3. Puesto que f es convexa, del teorema 6.2.1 y de 6.4.3 se deduce

$$f(\vec{x}) \geq f(\vec{x}^*) + Df(\vec{x}^*)(\vec{x} - \vec{x}^*) \geq f(\vec{x}^*) \quad \text{para todo} \quad \vec{x} \in D$$

que es lo que queriamos probar.

Teorema 6.4.4: Sea f una función diferenciable definida en un abierto A de un e.v.n.$\vec{E}$ con valores en $I\!R$. Una condición necesaria para que $\vec{x}^* \in A$ sea un mínimo local de f en A es que

$$Df(\vec{x}^*) = 0 \qquad 6.4.5$$

donde 0 representa la función nula en $\mathcal{L}(\vec{E}, I\!R)$.

Demostración: Sea $\vec{x}^* \in A$ un mínimo local de f en A. Puesto que A es un conjunto abierto, existe $B(\vec{x}^*, \varepsilon) \subset A$ tal que

$$f(\vec{x}^*) \leq f(\vec{x}) \quad \text{para todo} \quad \vec{x} \in B(\vec{x}^*, \varepsilon)$$

y como $B(x^*, \varepsilon)$ es un conjunto convexo, usando el teorema 6.4.2, concluimos que

$$Df(\vec{x}^*)(\vec{x} - \vec{x}^*) \geq 0 \quad \text{para todo} \quad \vec{x} \in B(\vec{x}^*, \varepsilon)$$

Haciendo el cambio de variable $\vec{v} = \vec{x} - \vec{x}^*$ en la relación anterior, vemos que ella es equivalente a

$$Df(\vec{x}^*)(\vec{v}) \geq 0 \quad \text{para todo} \quad \vec{v} \in B(\vec{0}, \varepsilon) \qquad (*)$$

lo que solo puede ocurrir si $Df(\vec{x}^*)$ es la función nula. En efecto, si existiera $\vec{z} \in \vec{E}$ tal que $Df(\vec{x}^*)(\vec{z}) > 0$, definiendo $\vec{v} := -\varepsilon \frac{\vec{z}}{\|\vec{z}\|}$ tendríamos que $\vec{v} \in B(\vec{0}, \varepsilon)$ y

$$Df(\vec{x}^*)(\vec{v}) = -\frac{\varepsilon}{\|\vec{z}\|} Df(\vec{x}^*)(\vec{z}) < 0$$

lo que contradice la relación $(*)$.

Nota 6.4.7: De acuerdo a la interpretación que se dió en la nota 5.1.1 a la cantidad $Df(\vec{x}^*; \vec{v}) = Df(\vec{x}^*)(\vec{v})$ (ver fórmula 5.2.5), vemos que la igualdad 6.4.5 corresponde al hecho que las pendientes de f en $\vec{x}^*$, en todas las direcciones, son nulas.

De acuerdo a la nota 5.2.2, la relación 6.4.5 equivale a decir que la aproximación lineal afín de f en $\vec{x}^*$ está dada por la función constante $h(\vec{x}) = f(\vec{x}^*)$. Geométricamente, 6.4.5 equivale a decir que en $\vec{E} \times I\!R$ el hiperplano afín tangente al grafo de f en $(\vec{x}^*, f(\vec{x}^*))$, es el hiperplano horizontal de ecuación $z = f(\vec{x}^*)$.

Nota 6.4.8: El mismo ejemplo de la nota 6.4.5 nos muestra que la relación 6.4.5 no es suficiente para que $\vec{x}^* \in A$ sea un mínimo local de f en A.
El próximo teorema nos muestra que para una clase particular de funciones esta condición es suficiente para que $\vec{x}^*$ sea un mínimo de f en el abierto A.

Nota 6.4.9: Es importante tener claro que el teorema anterior no es válido si $\vec{x}^*$ es un mínimo local de f en una parte convexa D de A. En efecto, si definimos la función $f : I\!R \to I\!R$ por $f(x) = x^2$ y el conjunto $D = [1, 2]$, es evidente que $x^* = 1$ es un mínimo de f en D y sin embargo $Df(1)(v) = 2v$ no es la función nula en $\mathcal{L}(I\!R, I\!R)$.

Nota 6.4.10: Si en el teorema anterior $\vec{E} = I\!R^n$ de acuerdo a la fórmula 5.2.13, la igualdad 6.4.5 es equivalente a

$$\nabla f(\vec{x}) = \vec{0} \qquad 6.4.6$$

Teorema 6.4.5: Sea f una función convexa y diferenciable definida en un abierto convexo Q de un e.v.n.$\vec{E}$ con valores en $I\!R$. Una condición necesaria y suficiente para que $\vec{x}^* \in Q$ sea mínimo de f en Q es que se tenga la relación 6.4.5.

Demostración: i) Del teorema anterior, aplicando el teorema 6.4.3 con $D = Q$, deducimos que la condición es necesaria.
ii) Supongamos ahora que $\vec{x}^* \in Q$ verifica 6.4.5. Puesto que f es convexa, del teorema 6.2.1 y de 6.4.5 se deduce

$$f(\vec{x}) \geq f(\vec{x}^*) + Df(\vec{x}^*)(\vec{x} - \vec{x}^*) = f(\vec{x}^*) \quad \text{para todo} \quad \vec{x} \in Q$$

que es lo que queriamos probar.

Teorema 6.4.6: Sea f una función de clase $\mathcal{C}^2$ definida en un abierto A de $I\!R^n$ con valores en $I\!R$. Una condición necesaria para que $\vec{x}^* \in A$ sea un mínimo local de f en A es que el Hessiano $H(\vec{x}^*)$ sea definido no negativo.

Demostración: Sea $\vec{x}^*$ un mínimo de f en A. Como A es abierto, existe $B(\vec{x}^*, \varepsilon) \subset A$. Del teorema 5.9.3 y la definición 5.9.2 deducimos que para todo $\vec{\delta} \in B(\vec{0}, \varepsilon)$ se tiene

$$f(\vec{x}^* + \vec{\delta}) = f(\vec{x}^*) + \langle \nabla f(\vec{x}^*), \delta \rangle + \frac{1}{2} \vec{\delta} H(\vec{x}^*) \vec{\delta}^t + o^2(\vec{\delta})$$

y como f es convexa, aplicando la fórmula 6.2.2, obtenemos

$$\vec{\delta} H(\vec{x}^*) \vec{\delta}^t \geq -2o^2(\vec{\delta}) \quad \text{para todo} \quad \vec{\delta} \in B(\vec{0}, \varepsilon) \qquad (*)$$

de lo cual se desprende fácilmente que $H(\vec{x}^*)$ es definida no negativa.
En efecto, sea $\vec{d} \in \mathbb{R}^n$ diferente de $\vec{0}$, aplicando la desigualdad $(*)$ a $\vec{\delta} := \lambda \vec{d}$, donde $0 < |\lambda| \leq \frac{\varepsilon}{||\vec{d}||}$ de modo que $\vec{\delta} \in B(0, \varepsilon)$, se deduce

$$\vec{d}H(\vec{x}^*)\vec{d}^t \geq -2||\vec{d}||^2 \frac{o^2(\lambda \vec{d})}{||\lambda \vec{d}||^2}$$

y tomando límite cuando λ tiende a 0, de la igualdad 5.9.3 obtenemos la desigualdad

$$\vec{d}H(\vec{x}^*)\vec{d}^t \geq 0$$

que es lo que queriamos demostrar.

Nota 6.4.11: El mismo ejemplo de la nota 6.4.5 nos muestra que la condición necesaria para mínimo local dada por el teorema anterior, no es una condición suficiente.

Nota 6.4.12: El mismo ejemplo de la nota 6.4.9 nos muestra que la condición necesaria para mínimo local dada por el teorema anterior, deja de serlo si $\vec{x}^*$ es un mínimo local de f en una parte D del abierto A.

Teorema 6.4.7: Sea f una función de clase $\mathcal{C}^2$ definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$. Una condición suficiente para que $\vec{x}^* \in A$ sea un mínimo local estricto de f en A es que $\nabla f(\vec{x}^*) = 0$ y $H(\vec{x}^*)$ sea definido positivo.

Demostración: i) Vamos a demostrar primero que si H es una matriz $n \times n$ definida positiva, entonces existe $\alpha > 0$ tal que

$$\vec{\delta}H\vec{\delta}^t \geq \alpha ||\vec{\delta}||^2 \quad \text{para todo} \quad \vec{\delta} \in \mathbb{R}^n \qquad (*)$$

Para esto, definimos la función $\phi : \mathbb{R}^n \to \mathbb{R}$ por $\phi(\vec{\delta}) = \vec{\delta}H\vec{\delta}^t$. Puesto que ϕ es una función continua en el compacto $c = \{\vec{\delta} \in \mathbb{R}^n : ||\vec{\delta}|| = 1\}$ y además $\phi(\vec{\delta}) > 0$ para todo $\vec{\delta} \in C$, del teorema 2.5.3 concluimos que existe $\alpha > 0$ tal que

$$\phi(\vec{\delta}) \geq \alpha \quad \text{para todo} \quad \vec{\delta} \in C$$

lo que es equivalente a $(*)$ puesto que para todo $\vec{\delta} \in \mathbb{R}^n$ se tiene que $\frac{\vec{\delta}}{||\vec{\delta}||} \in C$ y $\phi(\frac{\vec{\delta}}{||\vec{\delta}||}) = \frac{\phi(\vec{\delta})}{||\vec{\delta}||^2}$.
ii) Tal como decíamos en la demostración del teorema anterior, como A es abierto existe $B(\vec{x}^*, \varepsilon) \subset A$ y, del teorema 5.9.3 y la definición 5.9.2 deducimos que para todo $\vec{\delta} \in B(\vec{0}, \varepsilon)$

$$f(\vec{x}^* + \vec{\delta}) = f(\vec{x}^*) + \langle \nabla f(\vec{x}^*), \vec{\delta} \rangle + \frac{1}{2}\vec{\delta}H(\vec{x}^*)\vec{\delta}^t + o^2(\vec{\delta})$$

Aplicando ahora a la matriz $H(\vec{x}^*)$ lo que demostramos en i) y del hecho que $\nabla f(\vec{x}^*) = \vec{0}$, vemos que existirá $\alpha > 0$ tal que

$$f(\vec{x}^* + \vec{\delta}) \geq f(\vec{x}^*) + \alpha ||\vec{\delta}||^2 + o^2(\vec{\delta}) \quad \text{para todo} \quad \vec{\delta} \in B(\vec{0}, \varepsilon)$$

Para terminar la demostración bastará con mostrar que existe $\varepsilon' \in]0, \varepsilon]$ tal que
$\alpha ||\vec{\delta}||^2 + o^2(\vec{\delta}) > 0$ para todo $\vec{\delta} \in B(\vec{0}, \varepsilon')$ diferente de $\vec{0}$. Si no fuera así, existiría una sucesión $\vec{\delta}_n$ convergente

a $\vec{0}$, con $\vec{\delta}_n \neq \vec{0}$ para todo n, tal que $\alpha\|\vec{\delta}_n\|^2 + o^2(\vec{\delta}_n) \leq 0$ para todo n; dividiendo esta desigualdad por $\|\vec{\delta}\|^2$ y tomando límite sobre n obtenemos que $\alpha \leq 0$, lo que muestra una contradicción.

Nota 6.4.13: La condición suficiente para mínimo local estricto dada por el teorema anterior, no es una condición necesaria como lo muestra la función $f : I\!R \to I\!R$ definida por $f(x) = x^4$ que tiene como mínimo local estricto a $\vec{x}^* = 0$ y en ese punto el Hessiano no es definido positivo, en efecto $H(0) = 0$.

Nota 6.4.14: De las notas 6.4.1 y 6.4.2 deducimos fácilmente que los siete teoremas que hemos visto en esta sección se pueden enunciar cambiando el término "mínimo" por "máximo", el termino "función convexa" por "función cóncava" y el sentido de las desigualdades (a este último tipo de cambio corresponde el reemplazo del término "definido no negativo" por "definido no positivo" y "definido positivo" por "definido negativo" en los teoremas 6.4.6 y 6.4.7 respectivamente).

6.5 Mínimos con restricciones de tipo desigualdad. Teorema de Kuhn-Tucker

Nota 6.5.1: Los seis teoremas que en la sección anterior establecen condiciones necesarias y/o suficientes para mínimo local, se pueden dividir en dos grupos, los dos primeros (teoremas 6.4.2 y 6.4.3) y los cuatro últimos (teoremas 6.4.4 a 6.4.7). En los dos primeros, el mínimo o el mínimo local en cuestión está restringido a un subconjunto D del dominio de la función f. En los restantes, se trata de un mínimo o mínimo local no restringido, es decir en todo el dominio de la función f.
En esta sección vamos a dar una condición necesaria (y suficiente si la función es convexa) para mínimo local restringido de una función cuando el conjunto D toma una forma particular de gran importancia en el estudio de modelos matemáticos de la física y de la ingeniería.

Problema: Dadas n funciones convexas diferenciables $f_1, ..., f_n$, definidas en un abierto convexo Q de un e.v.n. $\vec{E}$ con valores en $I\!R$, vamos a estudiar los mínimos y mínimos locales de una función diferenciable $f_o : Q \to I\!R$, en el conjunto D definido por

$$D := \{\vec{x} \in Q : f_1(\vec{x}) \leq 0, ..., f_n(\vec{x}) \leq 0\} \tag{6.5.1}$$

Las funciones $f_1 ... f_n$ se llaman restricciones del problema y f_0 se llama función objetivo.

Definición 6.5.1: Si D es el conjunto definido en 6.5.1, dado $\vec{x}^* \in D$ definimos en $\vec{E}$ los conjuntos

$$\mathcal{D}(\vec{x}^*) := \{\vec{v} \in \vec{E} : \exists \mu > 0 \quad \text{tal que} \quad f_i(\vec{x}^* + \mu\vec{v}) \leq 0 \quad \text{para todo} \quad i : 1, ..., n\} \tag{6.5.2}$$

$$\mathcal{T}(\vec{x}^*) := \{\vec{v} \in \vec{E} : Df_i(\vec{x}^*)(\vec{v}) \leq 0 \quad \text{para todo} \quad i \in I(\vec{x}^*)\} \tag{6.5.3}$$

donde

$$I(\vec{x}^*) = \{i : f_i(\vec{x}^*) = 0\}. \tag{6.5.4}$$

Diremos entonces que las restricciones del problema en cuestión verifica la hipótesis de calificación en $\vec{x}^* \in D$, si se tiene la igualdad

$$\overline{D(\vec{x}^*)} = \mathcal{T}(\vec{x}^*) \tag{6.5.5}$$

Nota 6.5.2: En los teoremas 6.5.2 y 6.5.3 daremos condiciones suficientes bastante simples para que nuestro problema verifique la hipótesis de calificación de las restricciones en un punto $\vec{x}^* \in D$. Para esto necesitaremos tres lemas previos que damos a continuación.

Lema 6.5.1: El conjunto D definido por 6.5.1 es convexo.

Demostración: $\vec{a}, \vec{b} \in D \Rightarrow f_i(\vec{a}) \leq 0$ y $f_i(\vec{b}) \leq 0$ para todo $i : 1, ..., n$
$\Rightarrow \lambda f_i(\vec{a}) \leq 0$ y $(1-\lambda) f_i(\vec{b}) \leq 0$ para todo $\lambda \in [0,1]$ y todo $i : 1...n$
$\Rightarrow f_i(\lambda\vec{a} + (1-\lambda)\vec{b}) \leq 0$ para todo $\lambda \in [0,1]$ y todo $i : 1...n$
$\Rightarrow \lambda\vec{a} + (1-\lambda\vec{b}) \in D$ para todo $\lambda \in [0,1]$.

Lema 6.5.2: Si f es una función convexa definida en un convexo Q de un e.v.n. $\vec{E}$ con valores en $I\!R$ y si $\vec{x}^* \in Q$ y $\vec{v} \in \vec{E}$ verifican las desigualdades $f(\vec{x}^*) \leq 0$ y $f(\vec{x}^* + \mu\vec{v}) \leq 0$ para algún $\mu > 0$, entonces

$$f(\vec{x}^* + \lambda\vec{v}) \leq 0 \quad \text{para todo} \quad \lambda \in [0, \mu] \tag{6.5.6}$$

Demostración: Del teorema anterior sabemos que el conjunto $D = \{\vec{x} \in Q : f(\vec{x}) \leq 0\}$ es convexo. Para concluir basta ver que $\vec{x}^* + \lambda\vec{v} = \frac{\lambda}{\mu}\vec{x}^* + (1 - \frac{\lambda}{\mu})(\vec{x}^* + \mu\vec{v}) \in D$.

Lema 6.5.3: Si f es una función diferenciable definida en un abierto A de un e.v.n. $\vec{E}$ con valores en $I\!R$, y si $\vec{x}^* \in A$ y $\vec{v} \in \vec{E}$ verifican la desigualdad $Df(\vec{x}^*)(\vec{v}) < 0$, entonces existe $\mu > 0$ tal que

$$f(x^* + \lambda\vec{v}) < f(\vec{x}^*) \quad \text{para todo} \quad \lambda \in]0, \mu] \tag{6.5.7}$$

Demostración: Denotemos $r := Df(\vec{x}^*)(\vec{v}) < 0$. De la fórmula 5.2.5 y de la definición de límite vemos que existe $\mu > 0$ tal que

$$\frac{f(\vec{x}^* + \lambda\vec{v}) - f(\vec{x}^*)}{\lambda} \leq \frac{r}{2} \quad \text{para todo } \lambda \in]0, \mu]$$

Puesto que $\frac{r}{2} < 0$, multiplicando por λ obtenemos la desigualdad 6.5.7.

Teorema 6.5.1: Los conjuntos $\mathcal{D}(\vec{x}^*)$ y $\mathcal{T}(\vec{x}^*)$ definidos por 6.5.2 y 6.5.3 respectivamente verifican la inclusión

$$\overline{\mathcal{D}(\vec{x}^*)} \subset \mathcal{T}(x^*) \tag{6.5.8}$$

Demostración: Puesto que $\mathcal{T}(\vec{x}^*)$ es un conjunto cerrado, será suficiente demostrar que $\mathcal{D}(\vec{x}^*) \subset \mathcal{T}(\vec{x}^*)$. Sea $\vec{v} \in \mathcal{D}(\vec{x}^*)$ y $\mu > 0$ tal que $f_i(\vec{x}^* + \mu\vec{v}) \leq 0$ para todo $i \in I(\vec{x}^*)$. Del lema 6.5.2 se tiene para cada función f_i con $i \in I(\vec{x}^*)$ la relación 6.5.6, entonces dividiendo por $\lambda \in]0, 1]$ y tomando límite cuando $\lambda \to 0$, de la fórmula 5.2.5 obtenemos que $Df_i(\vec{x}^*)(\vec{v}) \leq 0$ para todo $i \in I(\vec{x}^*)$. Esto nos muestra que $\vec{v} \in \mathcal{T}(\vec{x}^*)$.

Teorema 6.5.2: Una condición suficiente para que las n restricciones de nuestro problema verifiquen la llamada hipótesis de calificación en $\vec{x}^* \in D$, dada por la igualdad 6.5.5, es que exista $\vec{x}_0 \in \vec{E}$ tal que

$$f_i(\vec{x}_0) < 0 \quad \text{para todo} \quad i : 1...n \tag{6.5.9}$$

(la relación 6.5.9 se llama condición de Slater y, es equivalente a decir que el interior del conjunto D definido por 6.5.1 es no vacío).

Demostración: Si $I(\vec{x}^*) = \phi$, el resultado es evidente. Supondremos entonces $I(\vec{x}^*) \neq \phi$. Como la inclusión 6.5.8 se tiene siempre, sólo debemos demostrar la contraria.

Dado $\vec{v} \in \mathcal{T}(\vec{x}^*)$ probaremos que $\vec{v} \in \overline{\mathcal{D}(\vec{x}^*)}$. Dividiremos la demostración en tres partes.

(i) De la relación 6.5.9, usando el teorema 6.2.1 vemos que $0 > f_i(\vec{x}_0) \geq f_i(\vec{x}^*) + Df_i(\vec{x}^*)(\vec{x}_0 - \vec{x}^*)$ para todo $i : 1, ..., n$ y como $f_i(\vec{x}^*) = 0$ para todo $i \in I(\vec{x}^*)$, definiendo $\vec{v}_0 := \vec{x}_0 - \vec{x}^*$, concluimos que

$$Df_i(\vec{x}^*)(\vec{v}_0) < 0 \quad \text{para todo} \quad i \in I(\vec{x}^*).$$

(ii) Es fácil ver ahora que si definimos $\vec{v}_k := \vec{v} + \frac{1}{k}\vec{v}_0$ se tiene para todo $k \in I\!N$ y todo $i \in I(\vec{x}^*)$

$$\begin{aligned} Df_i(\vec{x}^*)(\vec{v}_k) &= Df(\vec{x}^*)(\vec{v} + \frac{1}{k}\vec{v}_0) \\ &= Df(\vec{x}^*)(\vec{v}) + \frac{1}{k}Df(\vec{x}^*)(\vec{v}_0) \\ &< 0 \qquad \text{para todo} \quad i \in I(\vec{x}^*) \end{aligned}$$

y, de acuerdo al lema 6.5.3, esto implica que para cada $k \in I\!N$ y cada $i \in I(\vec{x}^*)$ existe $u_{k,i} > 0$ tal que

$$f_i(\vec{x}^* + \lambda\vec{v}_k) < 0 \text{ para todo } \lambda \in]0, u_{k,i}] \qquad (A)$$

(iii) Por otra parte, si $i \notin I(\vec{x}^*)$ se tiene que $f_i(\vec{x}^*) < 0$ y, por continuidad de las funciones f_i en $\vec{x}^*$, deducimos que para cada $k \in I\!N$ y cada $i \notin I(\vec{x}^*)$ existe $\mu_{k,i} > 0$ tal que

$$f_i(\vec{x}^* + \lambda\vec{v}_k) < 0 \quad \text{para todo} \quad \lambda \in]0, \mu_{k,i}] \qquad (B)$$

Si para cada $k \in I\!N$ definimos $\mu_k := \min\limits_{i:1,\ldots,n} \mu_{k,i} > 0$, de las relaciones (A) y (B) vemos que

$$f_i(\vec{x}^* + \mu_k\vec{v}_k) < 0$$

lo cual muestra que $\vec{v}_k \in \mathcal{D}(\vec{x}^*)$ para todo $k \in I\!N$. Como $\vec{v} = \lim \vec{v}_k$, concluimos que $\vec{v} \in \overline{\mathcal{D}(\vec{x}^*)}$

Teorema 6.5.3: Cuando $\vec{E}$ es un espacio de Hilbert, una condición suficiente para que las n restricciones de nuestro problema verifiquen la hipótesis de calificación en $\vec{x}^* \in D$, dado por 6.5.5, es que $Df_i(\vec{x}^*)$ con $i \in I(\vec{x}^*)$ sean linealmente independientes.

Demostración: No es difícil demostrar que si $Df_i(\vec{x}^*)$ con $i \in I(\vec{x}^*)$ son linealmente independientes, entonces debe existir $\vec{v}_0 \in \vec{E}$ tal que $Df_i(\vec{x}^*)(\vec{v}_0) < 0$ para todo $i \in I(\vec{x}^*)$. La demostración sigue entonces como la del teorema anterior (partes ii) y iii)).

Teorema 6.5.4: (Karush-Kuhn-Tucker) Si las restricciones de nuestro problema verifican la hipótesis de calificación en $\vec{x}^* \in D$, dada por la igualdad 6.5.5, entonces una condición necesaria para que $\vec{x}^* \in D$ sea un mínimo local de f_0 en D es que existan escalares $\lambda_i \geq 0$ para $i \in I(\vec{x}^*)$, llamados multiplicadores de Karush-Kuhn-Tucker, tales que

$$Df_0(\vec{x}^*) + \sum_{i \in I(\vec{x}^*)} \lambda_i Df_i(\vec{x}^*) = 0 \qquad 6.5.10$$

Si f_0 es una función convexa en Q, entonces 6.5.10 es una condición suficiente para que $\vec{x}^* \in D$ sea mínimo de f_0 en D.

Demostración: i) Demostremos primero que cuando f_0 es convexa en Q, entonces 5.6.10 es una condición suficiente para que $\vec{x}^* \in D$ sea mínimo de f_0 en D. Para ésto, de acuerdo al teorema 6.4.3, bastará demostrar que $Df(\vec{x}^*)(\vec{x} - \vec{x}^*) \geq 0$ para todo $\vec{x} \in D$. Como todas las restricciones de nuestro problema son convexas, de acuerdo al teorema 6.2.1, la igualdad 6.5.10 implica para todo $\vec{x} \in D$.

$$\begin{aligned} Df_0(\vec{x}^*)(\vec{x} - \vec{x}^*) &= -\sum \lambda_i Df_i(\vec{x}^*)(\vec{x} - \vec{x}^*) \\ &\geq \sum \lambda_i (f_i(\vec{x}^*) - f_i(\vec{x})) \\ &= -\sum \lambda_i f_i(\vec{x}) \geq 0 \end{aligned}$$

que es lo que queríamos probar.

ii) Supongamos ahora que $\vec{x}^* \in D$ es un mínimo local de f_0 en D, donde se verifica la hipótesis de calificación de las restricciones. Del teorema 6.4.2 sabemos que

$$Df(\vec{x}^*)(\vec{x} - \vec{x}^*) \geq 0 \text{ para todo } \quad \vec{x} \in D \tag{A}$$

y no es difícil demostrar, usando el lema 6.5.3, que esto equivale a decir que

$$Df(\vec{x}^*)(\vec{v}) \geq 0 \quad \text{para todo} \quad \vec{v} \in \mathcal{D}(\vec{x}^*)$$

Usando entonces la hipótesis de calificación de las restricciones concluimos que (A) equivale a la inclusión

$$\{\vec{v} \in \vec{E} : -Df(\vec{x}^*)(\vec{v}) \leq 0\} \subset \{\vec{v} \in \vec{E} : Df_i(\vec{x}^*)(\vec{v}) \leq 0 \text{ para todo } i \in I(\vec{x}^*)\}$$

El teorema 4.6.3 (Lema de Ferkas) nos permite concluir.