
CAPÍTULO 6

CONVEXIDAD Y EXTREMOS DE FUNCIONES DIFERENCIABLES

6.1 Introducción

En el siguiente capítulo se introduce una clase muy importante de funciones definidas sobre un e.v.n. a valores en \mathbb{R} . Aquella de las funciones convexas (y concavas). Esta clase, que es más general que la de las funciones lineales a valores en \mathbb{R} , aparece notablemente en numerosos modelos de las matemáticas aplicadas.

En primer lugar, se definen las funciones convexas, y se dan diversas caracterizaciones de aquellas que son diferenciable. También se muestran caracterizaciones de funciones convexas de aquellas funciones que son de clase \mathcal{C}^2 . Luego, se presentan una serie de condiciones necesarias y suficientes para que un elemento del e.v.n. sea mínimo de una función convexa sobre un conjunto determinado. Estos mismos resultados son extendibles a la clase de funciones concavas. Finalmente, cerramos el capítulo con el resultado mas importante de esta parte. El Teorema de Kuhn-Tucker, el cual nos entrega una condición necesaria y suficiente para los mínimos de una función sobre un conjunto dado por desigualdades de funciones convexas.

6.2 Funciones Convexas

Definición 6.2.1. Una función f definida en una parte convexa (ver Definición 4.3.1) Q de un e.v. \vec{E} con valores en \mathbb{R} , se dirá convexa en Q si para todo $\vec{x}, \vec{y} \in Q$ se tiene la

6.2. FUNCIONES CONVEXAS

desigualdad

$$f(\lambda\vec{x} + (1 - \lambda)\vec{y}) \leq \lambda f(\vec{x}) + (1 - \lambda)f(\vec{y}) \quad \text{para todo } \lambda \in [0, 1]. \quad (6.2.1)$$

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 6.2.2. Se llama epigrafo de una función f definida en una parte A de un e.v. \vec{E} con valores en \mathbb{R} , al conjunto

$$\text{epi}(f) = \{(\vec{x}, z) \in A \times \mathbb{R} : f(\vec{x}) \leq z\}. \quad (6.2.2)$$

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

Nota 6.2.1. De acuerdo a la definición anterior, la desigualdad (6.2.1) es equivalente a

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

que escrito de otro modo toma la forma

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

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 $\text{epi}(f)$ en $\vec{E} \times \mathbb{R}$.

Teorema 6.2.1. Dadas dos funciones convexas (resp. estrictamente convexas) f, g definidas en un conjunto convexo Q de un e.v. \vec{E} con valores en \mathbb{R} y dado $t \in \mathbb{R}_+$, 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)

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

6.3. CARACTERIZACIÓN DE FUNCIONES CONVEXAS DIFERENCIABLES

(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 6.2.2. *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 \mathbb{R} , si para cada $\vec{x} \in Q$ el conjunto $\{g_t(\vec{x}) : t \in T\}$ es acotado superiormente, entonces la función $f : Q \rightarrow \mathbb{R}$ definida por*

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

es convexa.

Demostración. De acuerdo con la Nota 6.2.1 basta con demostrar que $\text{epi}(f)$ es un conjunto convexo en $\vec{E} \times \mathbb{R}$. Vamos a mostrar entonces que $\text{epi}(f) = \bigcap_{t \in T} \text{epi}(g_t)$ y, como la intersección de conjuntos convexas es siempre convexa, habremos concluido que $\text{epi}(f)$ es convexo. La igualdad en cuestión es una consecuencia inmediata de la cadena de implicaciones

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

□

6.3 Caracterización de funciones convexas diferenciables

Nota 6.3.1. En los dos teoremas que siguen daremos una caracterización importante de la convexidad (resp. estricta convexidad) para funciones diferenciables.

Teorema 6.3.1. *Sea f una función diferenciable definida en un abierto convexo Q de un e.v.n. \vec{E} con valores en \mathbb{R} . Una condición necesaria y suficiente para que f sea convexa (resp. estrictamente convexa) es que*

$$\begin{aligned} f(\vec{x}) &\geq f(\vec{y}) + Df(\vec{y})(\vec{x} - \vec{y}) \quad \text{para todo } \vec{x}, \vec{y} \in Q \\ (\text{resp. } f(\vec{x}) &> f(\vec{y}) + Df(\vec{y})(\vec{x} - \vec{y}) \quad \text{para todo } \vec{x}, \vec{y} \in Q \text{ con } \vec{x} \neq \vec{y}) \end{aligned} \tag{6.3.1}$$

6.3. CARACTERIZACIÓN DE FUNCIONES CONVEXAS DIFERENCIABLES

Demostración. 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 (6.2.1).

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

Dividiendo por λ y tomando límite cuando $\lambda \rightarrow 0$ 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 (5.3.5) es equivalente a (6.3.1).

Supongamos ahora que f verifica (6.3.1). 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 . \square

Nota 6.3.2. Si en el teorema anterior $\vec{E} = \mathbb{R}^n$, la fórmula (6.3.1) se escribe

$$\begin{aligned} f(\vec{x}) &\geq f(\vec{y}) + \langle \nabla f(\vec{y}), \vec{x} - \vec{y} \rangle \quad \text{para todo } \vec{x}, \vec{y} \in Q \\ (\text{resp. } f(\vec{x}) &> f(\vec{y}) + \langle \nabla f(\vec{y}), \vec{x} - \vec{y} \rangle \quad \text{para todo } \vec{x}, \vec{y} \in Q \text{ con } \vec{x} \neq \vec{y}). \end{aligned} \quad (6.3.2)$$

Nota 6.3.3. En la Nota 5.3.2 veíamos que la ecuación $z = h(\vec{x})$, donde $h : \vec{E} \rightarrow \vec{F}$ 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 $f : Q \rightarrow \mathbb{R}$ 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 6.3.2. Sea f una función diferenciable definida en un abierto convexo Q de un e.v.n. \vec{E} con valores en \mathbb{R} . Una condición necesaria y suficiente para que f sea convexa (resp. estrictamente convexa) es que

$$\begin{aligned} [Df(\vec{x}) - Df(\vec{y})](\vec{x} - \vec{y}) &\geq 0 \quad \text{para todo } \vec{x}, \vec{y} \in Q \\ (\text{resp. } [Df(\vec{x}) - Df(\vec{y})](\vec{x} - \vec{y}) &> 0 \quad \text{para todo } \vec{x}, \vec{y} \in Q \text{ con } \vec{x} \neq \vec{y}). \end{aligned} \quad (6.3.3)$$

6.3. CARACTERIZACIÓN DE FUNCIONES CONVEXAS DIFERENCIABLES

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

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

sumando y teniendo en cuenta la linealidad de $Df(\vec{x})$ y $Df(\vec{y})$ se obtiene la desigualdad (6.3.3).

Supongamos ahora que f verifica (6.3.3). Dados $\vec{x}, \vec{y} \in Q$ definamos la función $\phi : [0, 1] \rightarrow \mathbb{R}$ 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 (6.3.3) 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 (6.3.1). La función f será entonces convexa. \square

Nota 6.3.4. Si en el teorema anterior $\vec{E} = \mathbb{R}^n$, la fórmula (6.3.3) se escribe

$$\langle \nabla f(\vec{x}) - \nabla f(\vec{y}), \vec{x} - \vec{y} \rangle \geq 0 \quad \text{para todo } \vec{x}, \vec{y} \in Q \quad (6.3.4)$$

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

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 \mathbb{R}^n los denotamos por una columna, el producto de una matriz H de $n \times n$ con un vector $\vec{\delta}$ de \mathbb{R}^n lo escribimos $H\vec{\delta}$ y el producto escalar del vector $H\vec{\delta}$ con un vector \vec{u} en \mathbb{R}^n lo escribiremos indistintamente $\vec{u}^t H\vec{\delta}$ o bien $\langle H\vec{\delta}, \vec{u} \rangle$, donde \vec{u}^t representa el traspuesto del vector \vec{u} , es decir el vector \vec{u} escrito como fila.

Teorema 6.3.3. *Sea f una función de clase \mathcal{C}^2 definida en un abierto convexo Q de \mathbb{R}^n con valores en \mathbb{R} . Una condición necesaria y suficiente para que f sea convexa (resp.*

6.3. CARACTERIZACIÓN DE FUNCIONES CONVEXAS DIFERENCIABLES

estrictamente convexa) es que para todo $\vec{x} \in Q$ su Hessiano $H(\vec{x})$ sea definido no negativo (resp. definido positivo), es decir

$$\begin{aligned} \vec{\delta}^t H(\vec{x}) \vec{\delta} &\geq 0 \quad \text{para todo} \quad \vec{\delta} \in \mathbb{R}^n \\ (\text{ resp. } \vec{\delta}^t H(\vec{x}) \vec{\delta} &> 0 \quad \text{para todo} \quad \vec{\delta} \in \mathbb{R}^n \quad \text{con} \quad \vec{\delta} \neq \vec{0}). \end{aligned} \quad (6.3.5)$$

Demostración. Supongamos que f es convexa. Dados $\vec{x} \in Q$ y $\vec{\delta} \in \mathbb{R}^n$ ($\vec{\delta} \neq \vec{0}$), definamos la función $\varphi : [0, \varepsilon] \rightarrow \mathbb{R}$, 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 $\varphi'(\lambda) = \langle H(\vec{x} + \lambda \vec{\delta}) \vec{\delta}, \vec{\delta} \rangle$. La desigualdad (6.3.5) equivale entonces a demostrar que $\varphi'(0) \geq 0$. Puesto que

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

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 $\varphi'(0) \geq 0$.

Supongamos ahora que para todo $\vec{x} \in Q$ la matriz $H(\vec{x})$ es definida no negativa. Dados $\vec{x}, \vec{y} \in Q$, si definimos $\vec{\delta} := \vec{y} - \vec{x}$ y φ como en la parte anterior 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 $\varphi'(\lambda) = \langle H(\vec{x} + \lambda \vec{\delta}) \vec{\delta}, \vec{\delta} \rangle \geq 0$ para todo $\lambda \in]0, 1[$, concluimos que $\varphi(1) - \varphi(0) \geq 0$ lo que equivale exactamente a la desigualdad (6.3.3) del teorema anterior. La función f será entonces convexa. \square

Nota 6.3.5. Del teorema anterior vemos que una forma cuadrática q definida en \mathbb{R}^n con valores en \mathbb{R} , esto es $q(\vec{x}) := \alpha + \langle \vec{c}, \vec{x} \rangle + \vec{x}^t H \vec{x}$, es convexa si y solo si la matriz H es definida no negativa.

Nota 6.3.6. 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 \mathbb{R}^n con valores en \mathbb{R} , será convexa si y solo si en todo punto de Q la aproximación cuadrática de f es convexa.

6.4 Funciones Concavas

Definición 6.4.1. Una función f definida en una parte convexa Q de un e.v. \vec{E} con valores en \mathbb{R} , se dirá cóncava (resp. estrictamente cóncava) en Q si la función $-f$ es convexa (resp. estrictamente convexa) en Q .

Nota 6.4.1. 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 6.2.1 habrá que cambiar el término convexas por cóncavas; en el Teorema 6.2.2 habrá que cambiar el término convexas por cóncavas y el supremo se cambia por un ínfimo; en el Teorema 6.3.1 habrá que cambiar el término convexa por cóncava y el sentido de la desigualdad (6.3.1); en el Teorema 6.3.2 habrá que cambiar el término convexa por cóncava y el sentido de la desigualdad (6.3.3) y finalmente; en el Teorema 6.3.3 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 (6.3.5).

6.5 Mínimos y máximos de una función

Definición 6.5.1. Dada una función f definida en un conjunto A de un e.v.n. \vec{E} con valores en \mathbb{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

$$\begin{aligned} f(\vec{x}^*) &\leq f(\vec{x}) \quad \text{para todo } \vec{x} \in B(\vec{x}^*, \varepsilon) \cap D & (6.5.1) \\ (\text{ resp. } f(\vec{x}^*) &< f(\vec{x}) \quad \text{para todo } \vec{x} \in B(\vec{x}^*, \varepsilon) \cap D \quad \text{con } \vec{x} \neq \vec{x}^*). \end{aligned}$$

Definición 6.5.2. Si la desigualdad (6.5.1) 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 6.5.1. 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 6.5.3. Dada una función f definida en un conjunto A de un e.v.n. \vec{E} con valores en \mathbb{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

$$\begin{aligned} f(\vec{x}^*) &\leq f(\vec{x}) \quad \text{para todo } \vec{x} \in D & (6.5.2) \\ (\text{ resp. } f(\vec{x}^*) &< f(\vec{x}) \quad \text{para todo } \vec{x} \in D \quad \text{con } \vec{x} \neq \vec{x}^*). \end{aligned}$$

6.5. MÍNIMOS Y MÁXIMOS DE UNA FUNCIÓN

Definición 6.5.4. Si la desigualdad (6.5.2) 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 6.5.2. 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 6.5.3. 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 6.5.1. Sea f una función convexa definida en una parte convexa Q de un e.v.n. \vec{E} con valores en \mathbb{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

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

Para concluir debemos demostrar que la desigualdad anterior también se tiene para todo $\vec{x} \in D \setminus B(\vec{x}^*, \varepsilon)$. 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 queríamos demostrar. \square

Teorema 6.5.2. Sea f una función diferenciable definida en un abierto A de un e.v.n. \vec{E} con valores en \mathbb{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

$$Df(\vec{x}^*)(\vec{x} - \vec{x}^*) \geq 0 \quad \text{para todo} \quad \vec{x} \in D. \quad (6.5.3)$$

Demostración. Sea \vec{x}^* un mínimo local de f en D . Dado $\vec{x} \in D$ definamos la función $\phi : [0, 1] \rightarrow \mathbb{R}$ 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

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

lo que corresponde exactamente a la desigualdad (6.5.3). \square

6.5. MÍNIMOS Y MÁXIMOS DE UNA FUNCIÓN

Nota 6.5.4. De acuerdo a la interpretación que se dió en la Nota 5.2.1 a la cantidad $Df(\vec{x}^*; \vec{x} - \vec{x}^*) = Df(\vec{x}^*)(\vec{x} - \vec{x}^*)$ (ver fórmula (5.3.5)), vemos que la desigualdad (6.5.3) corresponde al hecho que las pendientes de f en \vec{x}^* , en todas las direcciones interiores al convexo D , son no negativas.

Nota 6.5.5. Es fácil ver que la condición (6.5.3) 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 $f : \mathbb{R} \rightarrow \mathbb{R}$ definida por $f(\vec{x}) = x^3$, que verifica (6.5.3) 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 6.5.6. Si en el teorema anterior $\vec{E} = \mathbb{R}^n$, la desigualdad (6.5.3) se escribe

$$\langle \nabla f(\vec{x}^*), \vec{x} - \vec{x}^* \rangle \geq 0 \quad \text{para todo} \quad \vec{x} \in D. \quad (6.5.4)$$

Teorema 6.5.3. *Sea f una función convexa y diferenciable definida en un abierto convexo Q de un e.v.n. \vec{E} con valores en \mathbb{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 (6.5.3).*

Demostración. Del teorema anterior y, aplicando el Teorema 6.5.1, deducimos que la condición es necesaria.

Supongamos ahora que $\vec{x}^* \in D$ verifica (6.5.3). Puesto que f es convexa, del Teorema 6.3.1 y de (6.5.3) 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 queríamos probar. \square

Teorema 6.5.4. *Sea f una función diferenciable definida en un abierto A de un e.v.n. \vec{E} con valores en \mathbb{R} . Una condición necesaria para que $\vec{x}^* \in A$ sea un mínimo local de f en A es que*

$$Df(\vec{x}^*) = 0 \quad (6.5.5)$$

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

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

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

6.5. MÍNIMOS Y MÁXIMOS DE UNA FUNCIÓN

y como $B(x^*, \varepsilon)$ es un conjunto convexo, usando el Teorema 6.5.2, concluimos que

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

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

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

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 (*). \square

Nota 6.5.7. De acuerdo a la interpretación que se dió en la Nota 5.2.1 a la cantidad $Df(\vec{x}^*; \vec{v}) = Df(\vec{x}^*)(\vec{v})$ (ver fórmula (5.3.5)), vemos que la igualdad (6.5.5) corresponde al hecho que las pendientes de f en \vec{x}^* , en todas las direcciones, son nulas.

De acuerdo a la Nota 5.3.2, la relación (6.5.5) 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, (6.5.5) equivale a decir que en $\vec{E} \times \mathbb{R}$ 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 6.5.8. El mismo ejemplo de la Nota 6.5.5 nos muestra que la relación (6.5.5) 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 6.5.9. 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 $f : \mathbb{R} \rightarrow \mathbb{R}$ 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 $\mathcal{L}(\mathbb{R}, \mathbb{R})$.

Nota 6.5.10. Si en el teorema anterior $\vec{E} = \mathbb{R}^n$ de acuerdo a la fórmula (5.3.13), la igualdad (6.5.5) es equivalente a

$$\nabla f(\vec{x}) = \vec{0}. \quad (6.5.6)$$

Teorema 6.5.5. *Sea f una función convexa y diferenciable definida en un abierto convexo Q de un e.v.n. \vec{E} con valores en \mathbb{R} . Una condición necesaria y suficiente para que $\vec{x}^* \in Q$ sea mínimo de f en Q es que se tenga la relación (6.5.5).*

6.5. MÍNIMOS Y MÁXIMOS DE UNA FUNCIÓN

Demostración. Del teorema anterior, aplicando el Teorema 6.5.3 con $D = Q$, deducimos que la condición es necesaria.

Supongamos ahora que $\vec{x}^* \in Q$ verifica (6.5.5). Puesto que f es convexa, del Teorema 6.3.1 y de (6.5.5) se deduce

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

que es lo que queríamos probar. \square

Teorema 6.5.6. *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 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

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

y como f es convexa, aplicando la fórmula (6.3.2), obtenemos

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

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 $\vec{\delta} \in B(\vec{0}, \varepsilon)$, se deduce

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

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

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

que es lo que queríamos demostrar. \square

Nota 6.5.11. El mismo ejemplo de la Nota 6.5.5 nos muestra que la condición necesaria para mínimo local dada por el teorema anterior, no es una condición suficiente.

6.5. MÍNIMOS Y MÁXIMOS DE UNA FUNCIÓN

Nota 6.5.12. El mismo ejemplo de la Nota 6.5.9 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 6.5.7. 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. Vamos a demostrar primero que si H es una matriz $n \times n$ definida positiva, entonces existe $\alpha > 0$ tal que

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

Para esto, definimos la función $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ por $\phi(\vec{\delta}) = \vec{\delta}^t H \vec{\delta}$. 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 $\vec{\delta} \in C$, concluimos que existe $\alpha > 0$ tal que

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

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 $\phi(\frac{\vec{\delta}}{\|\vec{\delta}\|}) = \frac{\phi(\vec{\delta})}{\|\vec{\delta}\|^2}$.

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 $\vec{\delta} \in B(\vec{0}, \varepsilon)$

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

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

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

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 $\|\vec{\delta}_n\|^2$ y tomando límite sobre n obtenemos que $\alpha \leq 0$, lo que muestra una contradicción. \square

Nota 6.5.13. La condición suficiente para mínimo local estricto dada por el teorema anterior, no es una condición necesaria como lo muestra la función $f : \mathbb{R} \rightarrow \mathbb{R}$ 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 $H(0) = 0$.

6.6. MÍNIMOS CON RESTRICCIONES DE TIPO DESIGUALDAD. TEOREMA DE KUHN-TUCKER

Nota 6.5.14. De las notas 6.5.1 y 6.5.2 deducimos fácilmente que los siete teoremas que hemos visto en esta sección se pueden enunciar cambiando el término “mínimo” por “máximo”, el término “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 6.5.6 y 6.5.7 respectivamente).

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

Nota 6.6.1. 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 6.5.2 y 6.5.3) y los cuatro últimos (teoremas 6.5.4 a 6.5.7). 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, \dots, f_n , definidas en un abierto convexo Q de un e.v.n. \vec{E} con valores en \mathbb{R} , vamos a estudiar los mínimos y mínimos locales de una función diferenciable $f_0 : Q \rightarrow \mathbb{R}$, en el conjunto D definido por

$$D := \{\vec{x} \in Q : f_1(\vec{x}) \leq 0, \dots, f_n(\vec{x}) \leq 0\}. \quad (6.6.1)$$

Las funciones f_1, \dots, f_n se llaman restricciones del problema y f_0 se llama función objetivo.

Definición 6.6.1. Si D es el conjunto definido en (6.6.1), dado $\vec{x}^* \in D$ definimos en \vec{E} los conjuntos

$$\mathcal{D}(\vec{x}^*) := \{\vec{v} \in \vec{E} : \exists \mu > 0 \quad \text{tal que} \quad f_i(\vec{x}^* + \mu \vec{v}) \leq 0 \quad \text{para todo} \quad i : 1, \dots, n\} \quad (6.6.2)$$

$$\mathcal{T}(\vec{x}^*) := \{\vec{v} \in \vec{E} : Df_i(\vec{x}^*)(\vec{v}) \leq 0 \quad \text{para todo} \quad i \in I(\vec{x}^*)\} \quad (6.6.3)$$

donde

$$I(\vec{x}^*) = \{i : f_i(\vec{x}^*) = 0\}. \quad (6.6.4)$$

6.6. MÍNIMOS CON RESTRICCIONES DE TIPO DESIGUALDAD. TEOREMA DE KUHN-TUCKER

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

$$\overline{D(\vec{x}^*)} = T(\vec{x}^*). \quad (6.6.5)$$

Nota 6.6.2. En los teoremas 6.6.2 y 6.6.3 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 6.6.1. *El conjunto D definido por (6.6.1) es convexo.*

$$\begin{aligned} &\text{Demostración. } \vec{a}, \vec{b} \in D \Rightarrow f_i(\vec{a}) \leq 0 \text{ y } f_i(\vec{b}) \leq 0 \text{ para todo } i = 1, \dots, n \\ &\Rightarrow \lambda f_i(\vec{a}) \leq 0 \text{ y } (1 - \lambda)f_i(\vec{b}) \leq 0 \text{ para todo } \lambda \in [0, 1] \text{ y todo } i = 1, \dots, n \\ &\Rightarrow f_i(\lambda \vec{a} + (1 - \lambda)\vec{b}) \leq 0 \text{ para todo } \lambda \in [0, 1] \text{ y todo } i = 1, \dots, n \\ &\Rightarrow \lambda \vec{a} + (1 - \lambda)\vec{b} \in D \text{ para todo } \lambda \in [0, 1]. \quad \square \end{aligned}$$

Lema 6.6.2. *Si f es una función convexa definida en un convexo Q de un e.v.n. \vec{E} con valores en \mathbb{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*

$$f(\vec{x}^* + \lambda \vec{v}) \leq 0 \quad \text{para todo } \lambda \in [0, \mu]. \quad (6.6.6)$$

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

Lema 6.6.3. *Si f es una función diferenciable definida en un abierto A de un e.v.n. \vec{E} con valores en \mathbb{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*

$$f(\vec{x}^* + \lambda \vec{v}) < f(\vec{x}^*) \quad \text{para todo } \lambda \in]0, \mu]. \quad (6.6.7)$$

Demostración. Denotemos $r := Df(\vec{x}^*)(\vec{v}) < 0$. De la fórmula (5.3.5) 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 (6.6.7). \square

6.6. MÍNIMOS CON RESTRICCIONES DE TIPO DESIGUALDAD. TEOREMA DE KUHN-TUCKER

Teorema 6.6.1. *Los conjuntos $\mathcal{D}(\vec{x}^*)$ y $\mathcal{T}(\vec{x}^*)$ definidos por (6.6.2) y (6.6.3) respectivamente verifican la inclusión*

$$\overline{\mathcal{D}(\vec{x}^*)} \subset \mathcal{T}(\vec{x}^*). \quad (6.6.8)$$

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 6.6.2 se tiene para cada función f_i con $i \in I(\vec{x}^*)$ la relación (6.6.6), entonces dividiendo por $\lambda \in]0, 1]$ y tomando límite cuando $\lambda \rightarrow 0$, de la fórmula (5.3.5) obtenemos que $Df_i(\vec{x}^*)(\vec{v}) \leq 0$ para todo $i \in I(\vec{x}^*)$. Esto nos muestra que $\vec{v} \in \mathcal{T}(\vec{x}^*)$. \square

Teorema 6.6.2. *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 (6.6.5), es que exista $\vec{x}_0 \in \vec{E}$ tal que*

$$f_i(\vec{x}_0) < 0 \quad \text{para todo } i = 1, \dots, n \quad (6.6.9)$$

(la relación (6.6.9) se llama condición de Slater y, es equivalente a decir que el interior del conjunto D definido por (6.6.1) es no vacío).

Demostración. Si $I(\vec{x}^*) = \emptyset$, el resultado es evidente. Supondremos entonces $I(\vec{x}^*) \neq \emptyset$. Como la inclusión (6.6.8) 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 (6.6.9), usando el Teorema 6.3.1 vemos que $0 > f_i(\vec{x}_0) \geq f_i(\vec{x}^*) + Df_i(\vec{x}^*)(\vec{x}_0 - \vec{x}^*)$ para todo $i = 1, \dots, n$ 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

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

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

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

y, de acuerdo al Lema 6.6.3, esto implica que para cada $k \in \mathbb{N}$ y cada $i \in I(\vec{x}^*)$ existe $u_{k,i} > 0$ tal que

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

6.6. MÍNIMOS CON RESTRICCIONES DE TIPO DESIGUALDAD. TEOREMA DE KUHN-TUCKER

(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 $k \in \mathbb{N}$ y cada $i \notin I(\vec{x}^*)$ existe $\mu_{k,i} > 0$ tal que

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

Si para cada $k \in \mathbb{N}$ definimos $\mu_k := \min_{i=1, \dots, n} \mu_{k,i} > 0$, 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 $k \in \mathbb{N}$. Como $\vec{v} = \lim \vec{v}_k$, concluimos que $\vec{v} \in \overline{\mathcal{D}(\vec{x}^*)}$.

□

Teorema 6.6.3. *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 (6.6.5), 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 6.6.4 (Karush-Kuhn-Tucker). *Si las restricciones de nuestro problema verifican la hipótesis de calificación en $\vec{x}^* \in D$, dada por la igualdad (6.6.5), 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*

$$Df_0(\vec{x}^*) + \sum_{i \in I(\vec{x}^*)} \lambda_i Df_i(\vec{x}^*) = 0. \quad (6.6.10)$$

Si f_0 es una función convexa en Q , entonces (6.6.10) es una condición suficiente para que $\vec{x}^ \in D$ sea mínimo de f_0 en D .*

Demostración. Demostremos primero que cuando f_0 es convexa en Q , entonces (6.6.10) es una condición suficiente para que $\vec{x}^* \in D$ sea mínimo de f_0 en D . Para ésto, de acuerdo al Teorema 6.5.3, 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 6.3.1, la igualdad (6.6.10) implica para todo $\vec{x} \in D$.

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

6.6. MÍNIMOS CON RESTRICCIONES DE TIPO DESIGUALDAD. TEOREMA DE KUHN-TUCKER

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 6.5.2 sabemos que

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

y no es difícil demostrar, usando el Lema 6.6.3, que esto equivale a decir que

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

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 4.7.3 (Lema de Farkas) nos permite concluir. \square