9. Problemas sin restricciones

En un programa de este tipo las variables pueden tomar cualquier valor. Además, existen programas con restricciones de igualdad equivalentes a programas sin restricciones.

Definición 9.1 *Se llaman programas sin restricciones a los problemas del tipo*

$$\text{mín}\, f(\vec{x}) \ \textit{ó} \ \ \text{máx}\, f(\vec{x}),$$

con $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$*, siendo* D *el dominio de la función.*

Nuestro objetivo actual es el cáculo de estos óptimos. Pero notamos que, la determinación de los óptimos a partir de las definiciones solo es viable cuando las funciones son sencillas. Por tanto, hay que buscar métodos alternativos.

En economía lo que interesa es la máximo o el mínimo global. Asegurar esta idea de globalidad sólo nos puede venir al añadir la idea de convexidad o concavidad al programa. Por tanto, en nuestro caso sólo caracterizaremos los óptimos locales y no globales salvo en progamas convexos (respectivamente cóncavos).

En la resolución de los problemas de optimización, el teorema de Weierstrass nos da condiciones para que exista la solución. Sin embargo, aunque su utilidad en la búsqueda de la solución sea baja, nos da una garantia de la existencia de solucciones globales.

9.1. Condiciones necesarias de primer orden para mínimos y máximos

Sea $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ diferenciable en D abierto. En primer lugar caracterizaremos los óptimos en $\vec{x}^* \in D \subseteq R^n$ geométricamente, es decir mediante el plano tangente a la función en $\vec{x}^*$.

La ecuación del plano tangente a f en $\vec{x}^*$ es:

$$z = f(x_1^*, x_2^*) + \left(\frac{\partial}{\partial x_1} f(x_1^*, x_2^*), \frac{\partial}{\partial x_2} f(x_1^*, x_2^*) \right) \begin{pmatrix} x_1 - x_1^* \\ x_2 - x_2^* \end{pmatrix},$$

escrito de otra forma

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

Si $\vec{x}^o \in D \subseteq R^n$ es un óptimo de f en D, entonces el plano tangente es horizontal, es decir

$$\frac{\partial}{\partial x_1} f(x_1^*, x_2^*) = 0 \text{ y } \frac{\partial}{\partial x_2} f(x_1^*, x_2^*) = 0,$$

luego $z = f(x_1^*, x_2^*)$.

Generalizando lo anterior podemos enunciar el siguiente teorema para la función $f : D \subset R^n \to R$, diferenciable en D. Vamos a dar condiciones necesarias y suficientes para que un punto del interior del dominio D sea óptimo local de la función f.

Si D es abierto estas condiciones pueden aplicarse a todos los puntos de D.

Condición necesaria de primer orden.

Teorema 9.2 *Sea $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ diferenciable en D abierto. Si $\vec{x}^*$ es un mínimo o un máximo local de f en D. Entonces, se cumple que $\nabla f(\vec{x}^*) = \vec{0}$.*

Dem:

Demostraremos el máximo local.

Suponemos que $\vec{x}^* \in D$ es un máximo local de f en D. entonces, $f(\vec{x}^*) \geq f(\vec{x})$ para todo $\vec{x} \in B_\delta(\vec{x}^*) \cap D$ con $\delta > 0$ suficientemente pequeño.

Al se f diferenciable, las derivadas parciales son

$\dfrac{\partial f}{\partial x_i} = \lim\limits_{h \to 0} \dfrac{f(\vec{x}^* + h\vec{e}_i) - f(\vec{x}^*)}{h}$ para $i = 1, \ldots, n$. Para calcular este límite hacemos

$\lim\limits_{h \to 0^+} \dfrac{f(\vec{x}^* + h\vec{e}_i) - f(\vec{x}^*)}{h} \qquad \dfrac{\text{numerador es negativo}}{\text{denominador es positivo}} \to$ el cociente es ≤ 0.

$\lim\limits_{h \to 0^-} \dfrac{f(\vec{x}^* + h\vec{e}_i) - f(\vec{x}^*)}{h} \qquad \dfrac{\text{numerador es negativo}}{\text{denominador es negativo}} \to$ el cociente es ≥ 0.

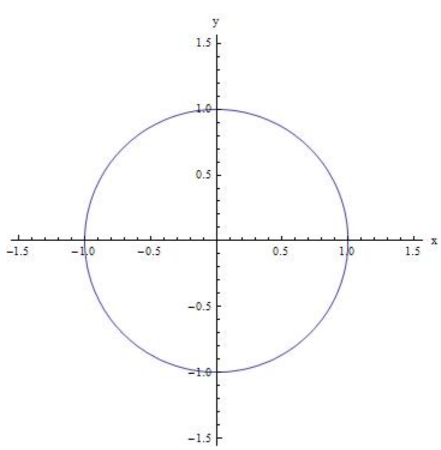
Como el límite existe, la única posibilidad es que sea 0. $\square$

Veamos un ejemplo en el cual es condición no es suficiente. Encontrar

$$\text{mín}(\text{máx}) x_1 x_2.$$

En $(x_1, x_2) = (0,0)$ tenemos que $\nabla f(0,0) = [0,0]$. En cambio, no existe ni máximo ni mínimo.

Si tomamos un entorno de radio $h > 0$ como el de la figura



tenemos que; en los puntos $(h,0)$, $(0,h)$, $(-h,0)$ y $(0,-h)$, la función $f(x_1,x_2) = 0$, pero, en los puntos (h,h) y $(-h,-h)$ la función es positiva y en los puntos $(-h,h)$ y $(h,-h)$ la función es negativa contradiciendo al afirmación de que es máximo o mínimo.

Observación 9.3 *Los puntos $\vec{x}^*$ que anulan el gradiente de una función f, se les denomina puntos críticos de la función f. En virtud de lo anterior, todos los máximos y mínimos de una función diferenciable son puntos críticos.*

Definición 9.4 *Al punto $\vec{x}^* \in F$ se le llama punto de silla de f en F si y sólo si*

a) El punto $\vec{x}^*$ es punto crítico de f en F.

b) Para todo $\delta > 0$ existen $\vec{x}^1$, $\vec{x}^2 \in B_\delta(\vec{x}^*) \cap F$ tales que $f\left(\vec{x}^1\right) < f(\vec{x}^*)$ y $f\left(\vec{x}^2\right) > f(\vec{x}^*)$.

9.2. Condiciones necesarias de segundo orden para mínimos y máximos

Si consideramos ahora una función $f : D \subset \mathbb{R}^n \to \mathbb{R}$ que tiene derivadas segundas continuas.

Teorema 9.5 *Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}$ que tiene derivadas segundas continuas.*

Si $\vec{x}^*$ es un mínimo local de f en D entonces, $Hf(\ \vec{x}^*)$ es semidefinida positiva.

Si $\vec{x}^*$ es un máximo local de f en D entonces, $Hf(\ \vec{x}^*)$ es semidefinida negativa.

Dem: Sólo el caso de mínimo.

Sea $\vec{x}^* \in D$ un mínimo local de f en D. Entonces, existe $r > 0$ tal que $f(\vec{x}^*) \leq f(\vec{x})$ para todo $\vec{x} \in B_r(\ \vec{x}^*) \cap D$.

Tomamos una dirección cualquiera $\vec{v} \in \mathbb{R}^n \setminus \left\{\vec{0}\right\}$ y queremos probar que $\vec{v}^T Hf(\ \vec{x}^*)\vec{v} \geq 0$.

Sea $\vec{x} = \vec{x}^* + \lambda\vec{v}$ y tomamos $\lambda = \dfrac{r}{\|\vec{v}\|}$, comprobamos que $\vec{x}$ pertenece a $B_r(\ \vec{x}^*) \cap D$.

$$\|\vec{x} - \vec{x}^*\| = \|\lambda\vec{v}\| = |\lambda|\,\|\vec{v}\| = \frac{r}{\|\vec{v}\|}\,\|\vec{v}\| = r.$$

Hacemos el desarrollo de Taylor de orden 3 en el intervalo $[\vec{x}, \vec{x}^*]$ obteniendo:

$f(\vec{x}) = f(\vec{x}^*+\lambda\vec{v}) = f(\vec{x}^*)+\nabla f(\vec{x}^*)\lambda\vec{v}+\frac{1}{2}\lambda^2\vec{v}^T Hf(\vec{x}^*)\vec{v}+\varepsilon\left(\vec{x}^* + \lambda\vec{v}, \vec{x}^*\right)\|\lambda\vec{v}\|^2$.

En virtud de que $\vec{x}^*$ es un mínimo local $(\nabla f(\vec{x}^*) = 0)$ tenemos que:

$f(\vec{x}) = f(\vec{x}^* + \lambda\vec{v}) = f(\vec{x}^*) + \frac{1}{2}\lambda^2\vec{v}^T Hf(\vec{x}^*)\vec{v} + \lambda^2\varepsilon\left(\vec{x}^* + \lambda\vec{v}, \vec{x}^*\right)\|\vec{v}\|^2$.

Como $\vec{x} \in B_r(\ \vec{x}^*) \cap D$ y $f(\vec{x}^*) \leq f(\vec{x})$, tenemos

$\frac{1}{2}\lambda^2\vec{v}^T Hf(\vec{x}^*)\vec{v} + \lambda^2\varepsilon\left(\vec{x}^* + \lambda\vec{v}, \vec{x}^*\right)\|\vec{v}\|^2 \geq 0$.

Dividiendo por λ^2 y tomando límites para $\lambda \to 0$, obtenemos

$\frac{1}{2}\vec{v}^T Hf(\vec{x}^*)\vec{v} + \lim_{\lambda\to 0}\varepsilon\left(\vec{x}^* + \lambda\vec{v}, \vec{x}^*\right)\|\vec{v}\|^2 \geq 0$, luego $\frac{1}{2}\vec{v}^T Hf(\vec{x}^*)\vec{v} \geq 0$ para todo $\vec{v}$.

Por tanto $Hf(\vec{x}^*)$ es semidefinida positiva. $\square$

Veamos un contraejemplo.

Ejemplo 9.6 *Comprobar, en el problema*

$$\text{mín } f(x_1, x_2) = x_1^3 + x_2^2,$$

que la implicación en sentido contrario del teorema 9.5 no es cierta.

9.3. Condiciones suficientes de segundo orden para mínimos y máximos

Si consideramos ahora una función $f : D \subset \mathbb{R}^n \to \mathbb{R}$ que tiene derivadas segundas continuas y un punto $\vec{x}^* \in D$ que es punto crítico cumpliendo $\nabla f(\vec{x}^*) = 0$. Buscamos condiciones que nos aseguren que $\vec{x}^*$ es un mínimo o un máximo local de f en D.

Lema 9.7 *Sea una función $f : D \subset \mathbb{R}^n \to \mathbb{R}$ que tiene derivadas segundas continuas.*
Si $Hf(\vec{x}^*)$ es definida positiva entonces, existe $r > 0$ tal que $Hf(\vec{x})$ es definida positiva para todo $\vec{x} \in B_r(\vec{x}^*)$.
Si $Hf(\vec{x}^*)$ es definida negativa entonces, existe $r > 0$ tal que $Hf(\vec{x})$ es definida negativa para todo $\vec{x} \in B_r(\vec{x}^*)$.

Teorema 9.8 *Sea una función $f : D \subset \mathbb{R}^n \to \mathbb{R}$ que tiene derivadas segundas continuas y un punto $\vec{x}^* \in D$ que es punto crítico cumpliendo $\nabla f(\vec{x}^*) = 0$.*
Si $Hf(\vec{x}^*)$ es definida positiva entonces, $\vec{x}^* \in D$ es un mínimo local de f en D.
Si $Hf(\vec{x}^*)$ es definida negativa entonces, $\vec{x}^* \in D$ es un máximo local de f en D.

En vista de los anteriores resultados y de la definición de punto de silla, obtenemos la siguiente caracterización del punto de silla.

Teorema 9.9 *Sea una función $f : D \subset \mathbb{R}^n \to \mathbb{R}$ que tiene derivadas segundas continuas y un punto $\vec{x}^* \in D$ que es punto crítico cumpliendo $\nabla f(\vec{x}^*) = 0$.*
Si $Hf(\vec{x}^*)$ es indefinida entonces, $\vec{x}^* \in D$ es un punto de silla de f en D.

Dem:
Si $\vec{x}^* \in D$ fuera un mínimo local, $Hf(\vec{x}^*)$ sería semidefinida positiva. Si $\vec{x}^* \in D$ fuera un máximo local, $Hf(\vec{x}^*)$ sería semidefinida negativa. Como $Hf(\vec{x}^*)$ es indefinida, no pude ser ni máximo ni mínimo. $\square$
Veamos un contraejemplo de que el recíproco del teorema anterior no es cierto.
Sea la función $f(x_1, x_2) = x_1^3 + x_2^2$.
Sabemos que $\nabla f(x_1, x_2) = [3x_1^2, 2x_2]$. En el punto $(0, 0)$ tenemos que $\nabla f(0, 0) = [0, 0]$.

Sabemos que $Hf(\ \vec{x}) = \begin{pmatrix} 6x_1 & 0 \\ 0 & 2 \end{pmatrix}$. En el punto $(0,0)$ tenemos que Hf $(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$.

Por el ejercicio , sabemos que $(0,0)$ es punto de silla y, por otra parte, Hf $(0,0)$ es semidefinida positiva.